

***LANDSCAPES OF STORIES: SEGMENTING AND ENCODING
CINEMATIC AND LITERARY NARRATIVES***

CATALINA IRICINSCHI, PhD

CORNELL UNIVERSITY 2015

COPYRIGHT PAGE

LANDSCAPES OF STORIES: SEGMENTING AND ENCODING CINEMATIC AND LITERARY NARRATIVES

Catalina Iricinschi, PhD

Cornell University 2015

ABSTRACT

Despite of a continuous perceptual experience of the environment, we discretize the world around us in units that can be processed and encoded for long-term use. These units, called events, are spatio-temporally bounded meaningful segments of our perceptual environment. We segment continuous streams of information by placing boundaries around consistent and coherent units very similar to the way we place boundaries around physical objects for future recognition. How exactly do we identify a boundary of a moving target such as the continuous stream of film, or our constantly changing lives, for that matter? Is the segmentation process a function of the nature of the stimulus, or is it an invariant across all instances of segmentation? In order to analyze event segmentation in dynamic stimuli, the studies to be presented here used film narrative as the default stimulus for the experimental tasks. Literary narrative was used to compare it with the more dynamic film.

In segmenting text and film, readers identify situation change boundaries in literary text and content-similar film adaptation. The linguistic nature of text elicits more temporal boundaries than the visual nature of film that, in turn, records a higher incidence of spatial boundaries. Breaking the causal and logical connections between adjacent space and time units makes a stimulus complex and difficult to encode. Segmentation is affected by the structural make-up of these complex stimuli. In parsing trailers for the identification of narrative acts (Exposition, Complication, and Resolution), the agreement across parsers is lower than the one reported in research for full-length feature films. Moreover, the Resolution segment is almost absent from the parsers segmentation thus verifying the promotional nature of the trailer: Holding back the most

appealing part of a narrative would lure the viewer into the movie theater. Puzzle films that scramble (albeit based on a strictly followed recipe) space and time continuity and contiguity result in a superficial parsing based mostly on low-level features of the films (e.g., change in color corresponding to change in levels for *Inception*, 2010).

Finally, space matters. A case study of segmenting cross-cultural films that present conflicts over space suggests that both the viewer's and the filmmaker's cultural background and biases affect film segmentation and processing.

BIOGRAPHICAL

Languages are fascinating beings. My undergraduate studies (and indeed many years before that) were dedicated to understanding the Spanish and English languages and the literatures they afforded. Three years of Linguistic studies at the graduate level at City University of New York preceded my studies in Psychology at Cornell University. From looking at the sentence as a structural unit of language, I moved on to researching the clusters of sentences that form a story. Narrative in both textual and cinematic format, together with the effort on the part of the reader and the viewer, became and remained my research interest.

ACKNOWLEDGEMENTS

First and foremost, I am genuinely grateful for the opportunity to study in the Psychology Department at Cornell University. The academic journey it afforded tested me and changed me in many ways; a list would lengthen the present acknowledgement page considerably. For all these tests and intellectual metamorphoses, I thank my professors and peers. For they allowed me to be curious and act on my curiosity, exercise it, practice it: a fundamental token of freedom for someone coming from a world where intellectual curiosity attracted punishment, almost in the religious sense of heresy and sin. And I did play the curiosity game, the best game I have ever played. Too good for victory to matter.

I learned from many. I know now what being knowledgeable really means. Thank you, James Cutting! With a little luck, a lot of work, and a long life, I may approximate a knowledgeable being sometime in the future. Many thanks to Shimon Edelman, Mike Goldstein, and David Field (among many others) for being always willing to share their knowledge with me. Pam Cunningham, Liz Chandler, and Cindy Durbin made it all happen when it was almost too late or almost impossible. My fellow graduate students made me feel fortunate: they are intelligent, supportive, generous, and witty. Jan Z., Ken Z., and Ralph T. have been true friends; they always caught me before I fell. I owe a big chunk of heartfelt gratitude to my friend forever and ever, Eduard Iricinschi: a true scholar, a remarkable human being.

TABLE OF CONTENTS

1. Segmenting, Parsing, and the Ensuing Events.....	7
1.1. Event Segmentation: A brief account.....	12
1.2. The Event Boundary: Segmentation in text and film narrative.....	17
1.3. Gaps in Narration: Segmenting versus Parsing and the Issue of Complexity.....	24
1.3.1. The Puzzle: Encoding structurally complex narrative plots.....	26
1.3.1.1. The structure of film trailers: Components of trailers and viewers' parsing strategies.....	27
1.3.1.2. Parsing Puzzle Film: Narrative content or perceptual salience?.....	55
2. Dimensions of Events: Space, Culture, and Stereotypes in Multi-Cultural Film....	62
2.1. Spatial Configurations and Narrative Maps.....	64
2.1.1. Patterns in narration: Views from above in visual media.....	66
2.2. Real Lands, Film Lands, Flatlands: Filming the Middle East; a case study.....	74
2.2.1. The Flatlands down below: Dialectics of camera angle in films of the Middle East; an <i>explanatory</i> case study.....	75
Concluding remarks.....	81
Bibliography.....	86
Filmography.....	93

1. Segmenting, Parsing, and the Ensuing Events

I look at the landscape, my gaze ranges over it, I see all sorts of distinct and indistinct movement; this impresses itself sharply on me, that is quite hazy. After all, how completely ragged what we see can appear! And now look at all that can be meant by “description of what is seen”.(Wittgenstein, c1997., pp. II, xi).

How completely ragged what we see can appear! I am commencing with Wittgenstein’s rather intriguing statement in an attempt to formulate the main concern that motivates the current research endeavors: A complex dynamic of continuous and discrete processes define our perception and encoding of the surrounding environment (and other continuous streams of information). The world around us, with its unrelenting motion and change, may provide our perceptual apparatus with a ‘ragged’ array of diverse stimuli. What emerges is, however, continuity in the context of perception. Firstly, perception itself is continuous in that there is no perceptual gap that we can report on (under normal circumstances), no split second of perceived absence of sensorial input. Secondly, despite the ‘raggedy’ input, the resulting impression is one of a stable, cohesive (and thus apparently continuous) world. But now, as Wittgenstein urges us, let us take a look at what the description of what is seen can suggest. If one is asked to describe what one sees in an outdoor environment, discrete well-organized bits would make the answer: trees, buildings, cars, and so on and so forth. A number of cognitive efforts support this answer – and I will elaborate on these efforts throughout the current work. Obviously, categorization as the process of identifying and clustering features that do *not* change from one instance to the next yields the list of types as classes of tokens. But there is a prerequisite for categorization and that is ‘cropping’ the

visual input along the correct lines in an effort to identify the cohesive units; in other words, we discretize the environment, impose boundaries onto a continuous domain. The bit of landscape that makes a sharp impression and the bit that is hazy are perceived as separate – “*this impresses itself sharply on me, that is quite hazy*”. In a sense, we make what we see appear ragged. But the *this* and the *that*, as segregated as they may be, are still integrated into the same smooth landscape. What we encode is by no means a mere collection of disparate bits; rather, we consolidate parts in unified wholes.

In sum, so far: Presented with stimuli that engage our perception continuously, we segment the input following laws of physics and perception, and we then restore the continuity by establishing causal (and otherwise) relationships between the segmented units.

Continuous streams of information are segmented into coherent units called *events*. From the domain of philosophy, to cognitive science, linguistics, neuroscience, and more recently, film cognition, event segmentation has been (or become) a topic of interest and research. The present work focuses on a particular application of event cognition: Event segmentation in narrative. Do we segment narratives into events during online processing in the same way we identify objects in a static image? Or are narrative events perceived as dynamic units that elicit the construction and update of contextual models? What features of the narrative determine and predict the event segmentation in narrative processing?

I refer to *ensuing* events in the title of this introductory note with the intention of distinguishing the stimulus (with its intrinsic properties) from the person performing the segmentation (with her experiential processing). The statistical analyses (presented later) of agreement across parsers have the underlying assumption that the stimulus is not (in most cases) over-specified such that it determines the segmentation outcome. Therefore, “[...] important elements and structure of an event are often, in some way, imposed by a person. There is presumably a reasonably high degree of uniformity across people in the way that they conceive of events. Thus, the components and structure of an event are not deterministically derived from the components of the world itself” (Radvansky & Zacks, 2014, p. 9).

Stories are man-crafted, they are cultural artifacts and, in this sense, they are not naturalistic stimuli. One may question the motivation behind studying event segmentation in narratives – Aren’t they already segmented, cropped, cut, structured, organized, and nicely trimmed? They are indeed, and this is precisely the advantage of investigating event segmentation in narratives. Composers of narrative, whether writers or filmmakers, edit and punctuate (i.e., insert breaks and boundaries in) their work with the intent of directing the audience’s attention and eliciting a certain impression of the narrative. But do readers and film viewers parse narratives along the pre-established event boundaries, the event markers intended by the author? Or do they part with the intended punctuation and re-construct their own movie? And, moreover, do readers and film viewers agree in their segmentation, or is narrative event segmentation a peculiar, experience-driven exercise that sets audience members apart? Moreover, the stimuli

used in collecting behavioral data tend to be simplified versions of events we encounter on a daily basis. These simplified versions are adequate in the controlled experimental environment, but they make for rather artificial stimuli. Clips from pre-edited feature film may be closer replicas of our daily environment and thus perceived as naturalistic stimuli.

In what follows, I will provide an account of the concept of *event* as it is understood in various domains of study as well as emphasize the event construals that apply to narrative processing (section 1.1). I will refer to existing research that investigates the concept of event in perception, language acquisition, and developmental psychology. Section 1.2 addresses the nature of the event boundary as a function of the stimulus. Narrative form spans a variety of media as composition languages. Do word-based and image-based narratives elicit events marked by different kinds of boundaries? Does event segmentation in language, in general, engage a different modality than segmentation in visual imagery? Theoretical and empirical evidence indicate a space-time distinction in the nature of event boundaries in film and novel, respectively. Although the result of editing and crafting, cinematic narrative is naturalistic in that it displays synchronous audio-visual stimuli. The editing and the crafting, however, can result in marked departures from what is considered naturalistic with respect to either temporal or spatial contiguity and contingency. Highly artificial and fairly bizarre, the film *Inception* (Nolan, 2010) serves as a good illustration of spatio-temporal discontinuity. Such compositions are complex and indeed puzzling (Botz-Bornstein, 2011; Buckland,

2009). Do we identify events even in these dis-continuous items? Section 1.3 will tackle the nature of event segmentation on cinematic narratives with structurally complex plots. The basic definition question, “what are events?” has been asked in many domains of study, and has built an impressive history especially in the realms of philosophy and psychology. Events are considered to be similar to objects in that they are spatio-temporally bounded units in both fields of study. “The simplest proposal as to what events are was made by Quine (1985/1996): Simply treat events as objects. That is, regard events as bounded regions of space-time. What we typically think of as objects (chairs and tables) are one family of bounded space-time regions, events are another” (Zacks & Tversky, 2001). The philosophical approach establishes various relationships between events and units of reasoning such as propositions or linguistic entities. “Events presumably are not linguistic entities; like trees and molecules, events can be talked about, referred to, and described but they are not themselves statements, sentences, descriptions, or any other kind of linguistic units. Nor are events propositions; propositions are supposed to be abstract entities, whereas events are spatio-temporally bounded particulars. Events and relations between events would exist even if there were no humans, or language, to describe them” (Kim, 1969).

The larger section 2 investigates space as construed in cinematic narrative. We encode the surrounding space reliably, we map spatial relationships of spatially adjacent items, we are endowed with a so-called spatial cognition. Can we then map narratives the way we map space? Can we represent narrative items and events in map-like configurations? Within a general discussion of abstract and physical space in narratives,

culture becomes a dimension of spatial representation. Analyses from a database on Middle Eastern film -- a culture engaged in a major conflict over issues related to space and physical borders – reveal inter-cultural patterns of cinematic representation.

The concluding remarks outline a unified view of the narrative parsing analyses presented in this thesis.

1.1. Event segmentation: A brief account

What is an event?¹ From informal approximations such as “the stuff that fills our lives” (J. B. Tversky, Zacks, & Hard, 2010, p. 216), or “what happens to us, what we do, what we anticipate with pleasure or dread, and what we remember with fondness or regret” (Radvansky and Zacks, 2014, p. 1), to more formal statements such as “a segment of time at a given location that is conceived by an observer to have a beginning and an end” (Zacks and Tversky, 2001), or -- in the context of perception -- “an organized whole, a whole built up of parts between which there are mutual influences” (Jansson, Bergström, Epstein, & Johansson, 1994), events have received an impressive array of definitions within many domains of study.

Philosophers have long concerned themselves with actions and events as basic ontological units and formal semanticists proposed logical constructions for event individuation. Although these fields provide relevant historical facts for the current

¹ “What is an event?” is the question with which Shipley and Zacks (2008) open their edited book on *Understanding Events*. And they continue: “An event may be miraculous, mysterious, seminal, even divine – and, of course, to paraphrase Ecclesiastes, there is one event that happeneth to us all. Of what do we speak when we say event?” (p. 3).

discussion on events, I will highlight only the definitions of events that apply to narrative processing.

In psycholinguistics, for instance, aspects of the relationship between language and mind are oftentimes inferred from studies on event processing. Cross-linguistic variation, for instance, delineates taxonomies of encoding motion verbs with some languages favoring verbs expressing the way in which motion is performed while other languages favoring verbs showing the trajectory of motion. The empirical interest here -- pertaining to the language and thought domain -- has to do with attention patterns displayed by speakers of the two language types while visually presented with motion events. Events received a more general definition in the field of psychology.

Various psychological studies on event processing are motivated by the overt intention of filling a specific theoretical or empirical gap in the research on events (e.g., Baggett, 1979, for segmentation in film and text; Magliano & Zacks, 2011, for exploring continuity editing in film; Swallow et al., 2010; Zacks & Swallow, 2007, for brain activity recording during event segmentation). While there are many aspects of event processing and event segmentation that have not been studied, the existing findings cover quite a bit of ground and reveal both the features of the event as a stimulus and the behavioral responses to those features.

Gestalt Psychologists advanced theories of coherent spatio-temporal clustering that generates events. In discussing the concept of order, Koffka (1935) refers to “an orderly march of events” as a sequence of coherently orchestrated actions: “An orderly march of events is, e.g., the movement of the piano keys when a practiced player plays a tune;

a mere sequence of events without any order takes place when the keys are pressed down by a dog running over the keyboard” (Koffka, 1935, p. 9). In Koffka’s account, sequences of occurrences are rendered orderly (non-random) by temporal contiguities that afford meaning and form sustained patterns. The functional isomorphism that Gestaltists advanced (e.g., Köhler, 1929) outlines spatial consistencies that, just like the above temporal constraints, define and confine an event.² Space and time appear to be, therefore, fundamental dimensions of events.

Johansson (1950) begins his analysis on event perception with a clear, very visual illustration. A birch tree in the wind will present the viewer with rather erratic small motions of the foliage and branches. “This is an example of unceasing motion, unceasing change” (p. 11). Despite the continuous and inconsistent change in the branches and leaves composing the structure of the viewed birch, the visual system does not fail to recognize the tree as a coherent unit. A picture of the birch, however, provides no motion, simply the static image of the tree. The perception of this image will be driven by the overall silhouette information and will extract a “relatively undifferentiated unit” (Johansson, 1950, p. 12). The unit perception, however, is not lost in the continuously moving exemplar because motion “has at the same time had an uniting and a segregating effect” (p. 12). Consistent with the Gestalt tradition, Johansson refers to an event as part event when motion (or other perceptual changes) affects only components of a whole unit. An event whole specifies unity even in the face of continuous change, if the change does not violate space-time coherence.

² Köhler’s functional isomorphism advocates mapping of external spatial structures onto brain areas. I am only referring to the aspect of encoding and maintaining spatial relationships across different instances as a strategy to extract patterns and identify spatially-bounded events.

The early studies of Heider and Simmel (1944) and Michotte's (1963) launching effect explain the concept of causality based on perceptions of spatio-temporal adjacencies in visual display of simple events. After viewing basic geometrical figures moving independently but consistent with spatiotemporal contiguities, participants reported rich narratives of causal (and even emotional) relationships between the figures (Heider and Simmel, 1944). The sequences of the motions of the geometrical figures were thus perceived as a coherent event due to the consistencies along the space and time dimensions.

In the domain of research on event segmentation strictly speaking, space and time may appear to be features that separate instead of uniting as they are the dimensions that afford the highest incidence of event boundaries. It is only logical, however, that events are segmented out based on their most salient features, the very features that keep them internally consistent.

Newtson (1976) provided the initial paradigm for researching event segmentation empirically. Participants were presented with videos of simple actions and asked to press a button whenever they thought an event ended and another was initiated. The resulting segmentation indicated that people agree significantly in their identification of event boundaries. Using the same paradigm, Zacks et al. (2001) show that events display a hierarchical organization with fine-grain subevents being contained in coarse-grain events. The fine-grain segmentation identified the smallest event unit whereas the coarse-grain segmentation marked the largest event. Newtson's procedure reveals important processes involved in online perception of continuous stimuli. The

methodology, however, calls the attention of the participant to the event boundary by providing explicit instructions regarding event boundary identification. Zacks et al's (2001) fMRI study by-passed this issue by asking the participants to watch passively the film stimuli and subsequently segment the same stimuli at fine and coarse grains. The observed patterns of brain activity were consistent with the occurrence of event boundaries displaying a gradual increase before the boundary and peaking immediately after the boundary. These findings indicate that information present in the stimulus affords the anticipation of event boundary.³ Discretizing the environment is a necessary effort (not to mention automatic in nature) because "the continuous input is so rich and complex that much of it must be, and is, ignored; the input must be categorized to be effectively processed and understood" (B. Tversky, Zacks, & Hard, 2008). And indeed, much of the continuous stimulus is ignored (or better said, forgotten). Research (e.g., Kurby & Zacks, 2008) shows that memory is more reliable in the vicinity of event boundaries. The Event Segmentation Theory (Kurby & Zacks, 2008; Zacks, Speer, Swallow, Braver, & Reynolds, 2007) explains the enhanced memory encoding at event boundary by claiming an updating and consolidation process of the current event at its end boundary. The empirical study of memory and event segmentation leads the discussion into the realm of segmenting narrative text and cinematic discourse as well as the nature of the event boundary. This body of research is discussed in section 1.2 below.

³ Such anticipation is part possible because of simplicity and familiarity of the activities shown in the stimuli. The feature film is a more complex stimulus, one that does not grant boundary anticipation easily. Moreover, contemporary films tend to be structurally complex and thus even more opaque to boundary anticipation.

1.2. The Event Boundary: Segmentation in text and film narrative

Literary and cinematic narratives afford different parsing strategies with respect to spatial-temporal dimensions. In a self-paced task, participants segmented excerpts from literary narratives and corresponding film adaptations into meaningful events. Parsing in the literary medium favors temporal event boundaries, whereas film segmentation tends to be driven by spatial shifts.

Research in text processing suggests that spatial and temporal event boundaries affect memory encoding (e.g., Bower & Rinck, 2001; Glenberg, Meyer, & Lindem, 1987; Morrow, Bower, & Greenspan, 1989; Rinck & Bower, 2000). If space and time distance is stipulated between events in the written stimulus, memory tends to decrease. This finding is consistent with Radvansky and Copeland's (2006) study indicating that doors in directly experienced places act in the same manner as the spatiotemporal distance in text: Walking through a door elicits situational update and consolidation of the event left behind; the door here functions as a spatial event boundary.

Baggett (1979) demonstrated that text and film stimuli that are structurally equivalent elicit almost identical recall. Even if the episodic structure of the two narrative media is not a complete overlap, it is reasonable to assume that "narrative itself is a deep structure quite independent of its medium". Despite the obvious constancies in the

two media, the written language and the moving image engage the audience differently. The nature of this distinction deserves a finer-grain approach (Chatman, 1980).

Magliano and Zack's (2011) fMRI study suggests that, in film, changes in action are more conspicuous event boundaries than spatial/temporal shifts. The continuity of the editing rules (Bordwell, 1985) seems to be more easily maintained along the spatial/temporal dimension than along the action/plot dimension. Two questions arise here. First, are spatial and temporal boundaries equally relevant to event segmentation or is one of these dimensions more responsible for the parsing outcome? Secondly, are spatial and temporal boundaries equally relevant to event segmentation *regardless the medium* or are they differently weighted in the two media?

Methods

Participants

Thirty graduate and undergraduate students at Cornell University participated in this narrative segmentation experiment. Three participants segmented each text and film excerpt used as stimuli. Two participants were excluded for not finishing the task and for providing a too fine parsing (with parsed units significantly smaller than the average short sentence length measured in words).

Materials and apparatus

Participants segmented excerpts of text and their corresponding film adaptation excerpts. The instruction simply required the identification of a change in situation thus eliciting the participants' "situation-change judgments" (Magliano, Miller, & Zwaan, 2001a). Prior to the participants' self-paced parsing, three adult raters compared the text

stimuli and their film adaptation excerpts and rated their similarity content-wise. Only the text-film stimuli pairs that were rated as highly similar were used in the experiment. In addition, an a priori analysis performed by myself and two research assistants marked all the temporal and spatial shifts in both the novel and the film stimuli. In identifying the temporal and spatial shifts in the stimuli used the situational model (e.g., Zwaan, Langston, & Graesser, 1995; Zwaan, Magliano, & Graesser, 1995; Zwaan, Radvansky, Hilliard, & Curiel, 1998) as a frame of work. Such shifts in text are mostly marked by evident temporal (e.g., *after*, *as soon as*) and spatial (e.g., *outside*, *to the city*) phrases respectively. For the film, however, the task of identifying the temporal and spatial markers independent of each other turned out to be challenging. If the film cuts to a different point in the narrative, it most likely cuts to a different location in the narrative as well. The instances of time shift in the absence of a space shift are extremely few. Shifts along the time dimension were identified using Magliano, Miller, & Zwaan's (2001) approach of identifying time shifts: Any sentence in the text and shot in the film presenting an action that is not simultaneous with or immediately following the preceding sentence or shot was assigned a value of one to code a time shift. Adjacent times expressed by adjacent sentence or shot units received a zero to indicate no time shift. The stimuli were therefore selected from five pairs of book-film adaptations: *Angela's Ashes*: Book–memoir 1996, Film – drama 1999; *Empire of the Sun*: Book–fiction/memoir 1984, Film–war film 1987; *House of Sand and Fog*: Book –drama 1999, Film–drama 2003; *Get Shorty*: Book–fiction 1990, Film–crime-comedy 1995; *Oliver Twist*: Book – fiction 1838, Film – 1948. The selection of the stimuli was also driven by

our goal of representing various genres and therefore we selected *Get Shorty* as a comedy, *Empire of the Sun* as a memoir, and *House of Sand and Fog* as drama.

The text segmentation was a paper-and-pencil self-paced test. Participants received the printed text and were instructed to mark on the text every change in situation. Moreover, they were asked to mark both the beginning and the end of situation changes (see Figure 1 below for an illustration).

GET SHORTY

/B/ When Chili first came to Miami Beach twelve years ago they were having one of their off-and-on cold winters: thirty-four degrees the day he met Tommy Carlo for lunch at Vesuvio's on South Collins and had his leather jacket ripped off. One his wife had given him for Christmas a year ago, before they moved down here. **/E/**

/B/ Chili and Tommy were both from Bay Ridge, Brooklyn, old buddies now in business together. Tommy Carlo was connected to a Brooklyn crew through his uncle, a guy named Momo, Tommy keeping his books and picking up betting slips till Momo sent him

Figure 1. Parsing text illustration: Participants marked the beginnings and endings of any situation change they identified.

The segmentation of the corresponding film adaptation was, by the inherent nature of the stimulus, a bit more complex. Participants watched the film excerpts on a split screen with half of the screen occupied by the running film and half by a spreadsheet. For every situation change, the viewer had to record the corresponding frame number on the spreadsheet. This may sound as more cumbersome than it actually was: The film stimuli were run on a video software that displays the frame number. The recording of the situation change took only a copy-and-paste operation on

the part of the participant. An illustration of the experimental setting is provided in Figure 2 below.

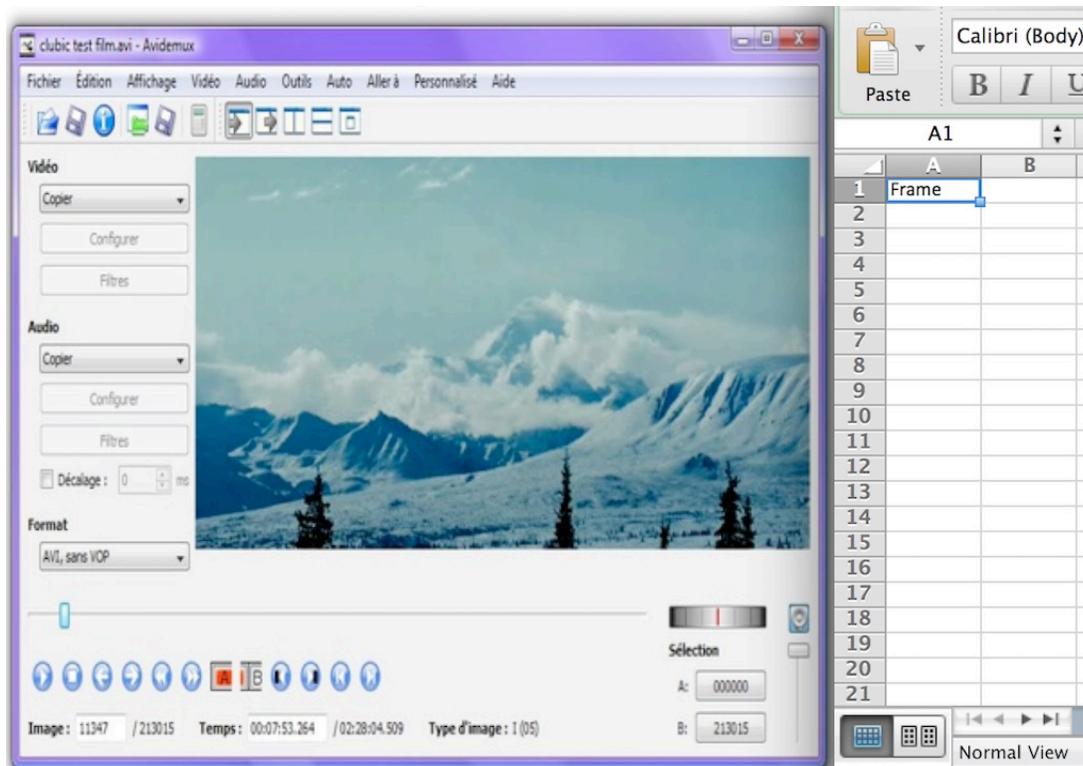


Figure 2. Film segmentation experimental setting.

Analyses and Results

The incidence of time and space shifts as identified by the participants was measured in both text and film. The percentage of space shifts identified by the parsers compared to the percentage of space shifts provided by the a priori analysis is higher in film than in text. In other words, a higher percentage of the space shift was recorded in film (mean = 9.38%) than in text (mean = 3.33%). An analysis of variance of the time and space shift occurrence indicates a significant difference between space shifts across media ($t=4.31$, $df=8$, $p<.005$). Although there is no significant difference between the occurrence

of time in the two media of text and film narrative, there is a trend favoring the speculation that text elicits more time shifts than the visual medium of film.

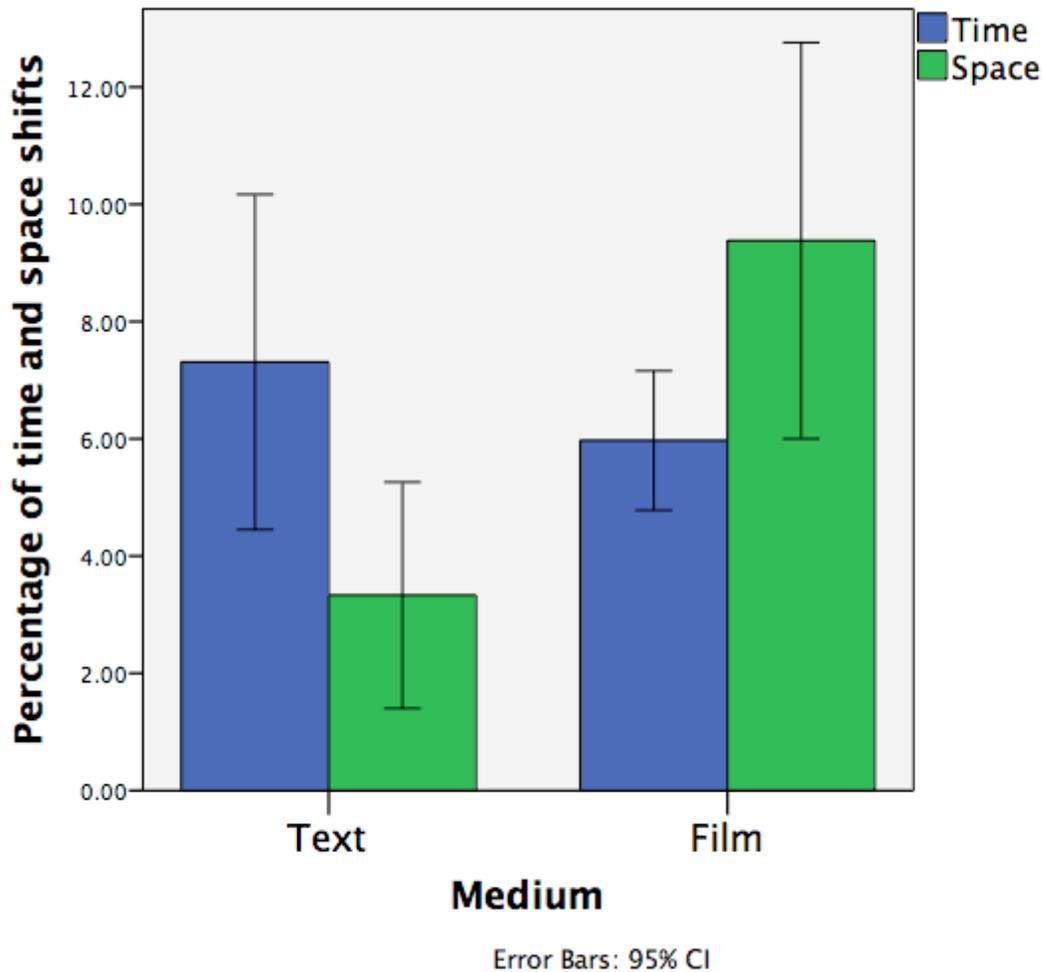


Figure 3. Percentage of time and space shifts identified by text readers and film viewers. The language medium of the literary text affords a higher incidence of time shifts than the visual medium of film. Space shifts, on the other hand, are perceived as being more frequent in film than time shifts.

When broken down per pair of items, space records consistently a lower incidence in text than in film independent of genre (see figure below).

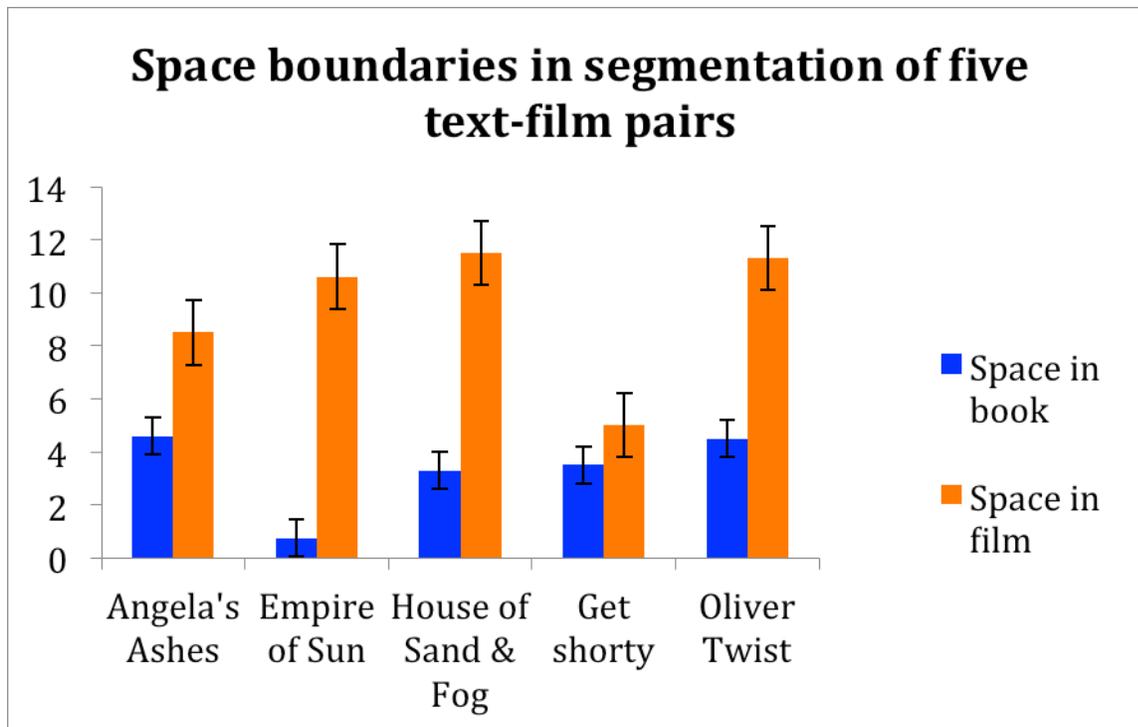


Figure 4. Space shifts in the situation-change segmentation participants performed on five text-film pairs. Literary narrative seems to elicit fewer spatial boundaries independent of genre.

The interobserver agreement value of $\kappa=0.43$ indicates that parsers moderately agree in their segment identification and segment temporal and spatial nature.

In conclusion, written stimuli elicited a higher percentage of temporal event boundaries than spatial event boundaries. A few aspects of the observed segmentations are worth noting. In their segmentation of text, participants left out of their segmentation the descriptive paragraphs and the paragraphs elaborating on characters' mental states. It is true that the text segmentation task required the identification of both the beginning and the ending of an event marked by a situation change. The film segmentation task

only elicited the transitions into a new change. Intrigued by the non-events found in the text segmentation, I informally asked two voluntary participants to perform the beginning and ending confounded segmentation task in film. No bit was left out. Apparently, Hitchcock's quote defining drama as life with the dull bits left out carries some truth.

With additional analyses to be carried out, the investigation of the text and film segmentation can be informative in a number of ways. It can contribute to the lengthy discussion (for a summary, see Elliott in Ryan, 2004) regarding the relationship between the literary narrative and its film adaptation: Do we conceive of the two media as one main form of discourse (literary) with a subordinate visual expression, or do we treat them as independent forms of narrative? A more important aspect for the current discussion has to do with the definition proposed in the research literature for an event. If events are perceived differently in different media, then the meaningful unit that we refer to as event should be operationally defined as contingent upon the form of expression. Why are descriptions of mental states and landscapes *not* events? It is possible that the notion of event may be conceived as too close to an action. If so, the experimental instructions given in the research lab should be worded with this in mind.

1.3. Gaps in Narration: Segmenting versus Parsing and the Issue of Complexity

In the introduction of the present account on event segmentation, events were defined as spatio-temporally bounded entities (Kim, 1969; Zacks & Tversky, 2001). With narrative being artifactual items, the space and the time dimensions are in the hands of the story maker. As Bordwell (1985) argues, narrative and narration are by no means

synonyms; they inform and modulate each other with narration providing one mode (of many possible) of viewing and experiencing the narrative. Now what if the narrative creator chooses to craft a mode of narration that breaks down the continuity and contiguity of time and space? What if a narrative unit that would be perceived as a unitary event is broken down and spread over disparate bits of time and space that do not obviously cohere? In other words, what temporal and spatial gaps are introduced in the fabric of narration such that logically adjacent bits are displaced far apart (in the diegetic time) from each other? Film viewing equated mere entertainment, a rather passive form of entertainment on the part of the viewer. Bordwell (1985), among many others, argues against the uninvolved film viewing account and places the viewer in an engaged dynamic of narrative processing. "The viewer must take as a central cognitive goal the construction of a more or less intelligible story. But what makes something a story? And what makes a story intelligible?" (Bordwell, 1985, p. 33). The question of narrative intelligibility becomes even more pressing in the recent film narration modes of scrambled temporal and spatial orders. How do viewers make sense of orders that defy logical and chronological connections? How do viewers make sense of the *Inception* (2010) characters that have the ability to inhabit multiple times and multiple spaces at the same time? Not only are these spaces and times disconnected by artifacts of narration, but in order for the story to be somewhat intelligible these times and spaces should not be far apart with large conceptual gaps between them. "There are, of course, degrees and kinds of chains, gaps, and reversals of space; and our recognition of the kinds will depend on the nature of other conventions governing, say, camera placement

[...]. Connecting screen spaces to a pattern of story space⁴ does not prohibit also using gaps or other distortions to create a story space which is *not* the sum of spatial fragments on the screen” (Branigan, 2013, p. 44). The viewer does indeed use gaps and distortions of space and time and may even enjoy it. Viewers of *Inception* (2010), for instance, report enjoying the film despite not being able to give a satisfactory account of the plot; at least not after the first viewing. It may be the case that the puzzle films engage the viewer in a learning process that requires multiple viewings. “Every film trains its spectator”, claims Bordwell (1985, p. 45). Do films like (*Inception*, c2010.) and (*Memento*, 2002) train their spectators? After how many viewings?

The section below analyzes the structure of complex narrative plots (un-bounded temporally and spatially) and the outcome of viewers’ event segmentation of these non-linear narrative plots.

1.3.1. The Puzzle: Encoding structurally complex narrative plots

The *puzzle plots* analyzed in this section come from films that violate temporal orders beyond the now habitual flashback and flashforward, or even time travels. Space in these film narratives is subordinate to the inconsequential narration. Two kinds of spatio-temporal complexity are discussed here. One is the film trailer: sequences of unfinished disparate bits of film meant to incite the viewer’s curiosity and motivate watching the entire film. The other comes from the puzzling film plot illustrated in (*Inception*, c2010.) and (*Memento*, 2002) due to their broken spatio-temporal order. These films, I argue, are structurally complex due to their lack of linearity in narration.

⁴ I will elaborate on screen space and story space with an illustration of multicultural film in section 2.

The puzzle plot breaks continuities and builds discontinuity by placing logically- and causally-linked narrative bits in long-distance dependencies. The analyses will address structural aspects of trailer and film samples together with viewers' segmentation data.

1.3.1.1. The structure of film trailers: Components of trailers and viewers' parsing strategies

Research on event cognition and event segmentation involves a vast array of stimuli and methodologies, from self-paced segmentation of text and film to computer interfaces recording online event boundary identification. The questions addressed in these empirical settings target (among other phenomena) agreement across participants, the nature of event boundary as a function of the stimulus, and memory and attention within and across events.

Film and video clips have been consistently used as stimuli in event segmentation experiments. The visual nature of the film clip stimuli allows for testing whether events are as robust as objects in participants' perception of 'edges' in dynamic and static environments, respectively. The clips used as experimental stimuli tend to be highly crafted and controlled in order to constrain the participants' behaviors and make them interpretable in relation to the research question addressed. Several studies investigating event segmentation used *The Red Balloon* (Lamorrisse, 1957) as the stimulus (e.g., Baggett, 1979, for segmentation in film and text; Magliano & Zacks, 2011, for exploring continuity editing in film; Swallow et al., 2010; Zacks & Swallow, 2007, for brain activity recording during event segmentation). While the absence of dialogue makes *The Red Balloon* a very appropriate stimulus for investigating the processing of

visual stimuli, the narrative in the film is causally tight. With only a couple exceptions (the grandmother, anxiously waiting for the boy, looks out the window, and the balloon chase scene), every shot of *The Red Balloon* features at least one of the two main characters, the boy and the red balloon. The narrative is thus anchored in the well-defined main characters while assuring the continuous editing across the sequential events the characters experience.

Generally speaking, the film narratives presented in these controlled experiments are therefore linear, where linearity is operationally defined as a strong causal relationship between any given inherent event marked intrinsically by formal edits already present in the stimulus and the following event.⁵

Film trailers, on the other hand, are rather messy stimuli, not particularly well behaved along the linearity dimension. Trailers present the viewer with snippets of the original story in a sequence that deliberately breaks down (or at least reduces) the causal link between events. Why then study event segmentation in film trailers? In what way can film trailer segmentation be informative? And more importantly, what is the motivation behind studying film trailer segmentation?

The research on event segmentation investigates the process by which participants segment continuous streams of information. The continuity here does not refer only to the nature of the uninterrupted presentation of the stimuli, but also to their content. As

⁵ By inherent event I denote the chunk of narrative contained within the edits characteristic to the medium, chapters and paragraphs for text and shots for film. The discussion on the cognitive implications of text edits (with the initial breathing-contingent punctuation) and film edits (with the initial one-shot films) could be lengthy and indeed engaging. For the purposes of the current consideration, however, the notion that the narrative media comes with built-in segments suffices.

mentioned above, the film stimuli tend to be linear stories easy to follow due to the apparent causal connections between adjacent bits. Trailers are nonlinear, discontinuous. While *Erin Brockovich* (“Erin Brockovich,” c2000.) maintains bits of the story line and hints at the happy ending, *Mission Impossible II* (2000) is a sequence of disparate shots rapidly moving from rock climbing to fight scenes and car chases while featuring (apparently randomly) sun glasses and fires. The information that a film trailer such as *Mission Impossible II* provides reliably is the genre of the film.

The trailer’s promotional intent attracted research interests from both filmmakers and marketers. The current general consensus sees the trailer as a new form of narrative and not only as a commercial devoid of narrative content (Kernan, 2009; Maier, 2009, 2011; Suckfüll & Moellering, 2015).

Film trailers, released with apparent promotional intent, are condensed versions of feature films. Despite being widely available in many versions and receiving substantial attention from consumers, film trailers are mostly overlooked in the domain of empirical research. Trailers have specific features that may define a distinct genre of visual media. Inherent to film trailers, the “complexity of editing strategies and special effects” (Kernan, 2004; Maier, 2009) indicates a focus on presentation and not on narrative coherence. Contrary to continuity editing techniques (Bordwell, 2012; Smith, 2012), trailers emphasize between-shot transitions by inserting conspicuous dissolves, fades, and black screens. The questions I am addressing have to do with the factors that affect event segmentation. What aspects of the stimulus may reduce or enhance (even if

superficially) the agreement among participants in their event segmentation? Do genre and year of production inform us on the nature of the stimulus and the segmentation outcome? And, most importantly, are trailers viewed as forms of narrative discourse, or are they simply fulfilling only a commercial role?

Film trailers use the feature film format to display diegetic features that elicit the viewer's engagement with the narrative as well as the non-diegetic features revealing information about the film and thus eliciting a more distant and critical stance in the viewer. By manipulating the diegetic and non-diegetic aspects, Suckfüll and Moellering (2015) created trailers that targeted viewers who are more sensitive either to the qualitative nature and emotional benefit of the narrative or to the formal features of the trailer. The different trailer versions thus created successfully segmented the audience based on the corresponding qualitative or production-related Modes of Reception (the authors' term). Trailers, therefore, can be narratives as well as forms of advertisement with just a thin narrative gloss. "While trailers are a form of advertisement, they are also a unique form of narrative exhibition, wherein promotional discourse and narrative pleasure are conjoined (whether happily or not)" (Kernan, 2009, 1). The more precise motivation behind the current study addresses the narrative structure of film trailers (as similar or dissimilar from their entire feature film counter parts) as well as the segmentation strategies viewers apply to this peculiar visual narrative discourse. In order to investigate these features, a more detailed discussion elaborating on structural aspects of trailers is in order.

An analysis of the structural features of the trailers in the database shows a shot duration pattern that replicates the gradually shorter shot in Hollywood film from 1930 until now (Cutting et al.,2011; Salt, 1983). With the exception of the 1950 trailers (average shot duration 5.63 seconds) being higher than the average shot duration in

1940 trailers (3.13 seconds), the trailers from 1960 to 2010 consistently decrease in shot duration. With dissolves and fades also decreasing the evolution of trailers mimic the evolution of Hollywood film. The decreasing trend of ASL in film over decades (with the above-mentioned exception of the 1950-1960 decade) seems to simply scale down in the case of trailers: While showing a similar ASL pattern, trailers are faster-paced due to shorter shot durations (see Figure 5 below).

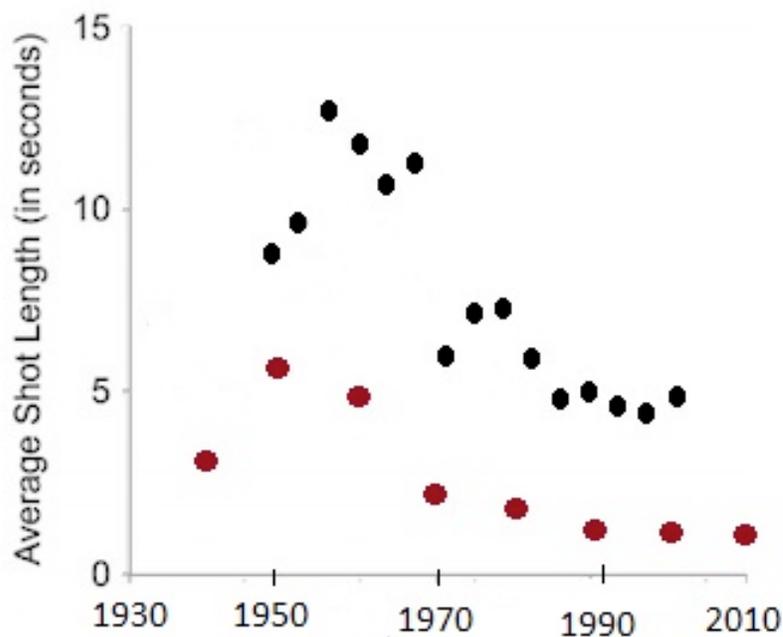


Figure 5. Similar trends of shot duration in film trailers and full-length films between 1930 and 2010. The red dots represent the group mean ASLs for 21 film trailers. The black dots represent the group mean ASLs for 150 films (adapted from Cutting, DeLong, and Brunick, 2010). The ASL in film trailers is consistently lower.

With continuity editing consistently violated, film trailers approximate the structural make-up of complex (or puzzle) narrative plots. Complexity in film has come under empirical scrutiny recently both with an interest in formal features of complex filmic

narration and with an interest in viewers' perception and comprehension of discontinuous times and spaces in puzzle plots. While the complex filmic narrative tends to sustain an underlying causality and logic, trailers deliberately shy away from plot-related causal relationships. Parsing a continuous stream of information has been correlated with measures of understanding and encoding information. How do viewers parse a trailer and extract a storyline given the lack of causality between shots? Is the trailer a narrative form?

The art of film and that of theater play are evidently related, with film implementing aspects of the older and more experienced dramatic work of the theater stage. Initially thought of as imitations of life, successful plays had to display a structured sequence of events that cohered to form a three-segment unit. In his *Poetics*, Aristotle asserts a three-segment tragedy plot as unitary and complete. "Now a thing is a whole if it has a beginning, a middle, and an end. [...] A well-constructed plot, therefore, will neither begin at some chance point nor end at some chance point [...]" (Aristotle, c1982).⁶ The film plot, at least from the perspective of the screenwriter, has maintained the three-act structure of the play in an approximately thirty-sixty-thirty-minute sequence of film segments. The three-act structure as seen by Field (Field, c1994.) is depicted in the figure below.

⁶ In Aristotle's view, a "unified imitation" as expressed by a mimetic art is rather rigid and tightly hinged. "[...] the parts of the plot must be so organized that if any one of them is displaced or taken away, the whole will be shaken and put out of joint". Modern film moved away from such strictures. Although "Hollywood favors unified narratives" with a rather rigid cause-effect sequence forming "an unbroken chain across the film", the film form accommodates the "dangling cause" defined as "information or action which leads to no effect or resolution until later in the film" (Thompson, 1999). Film trailers break even the "loose" structure of the new Hollywood full feature film by inserting frequent turning points, re-turning points (to earlier scenes), and *suspended* dangling causes the resolution of which lure the viewer into the movie theater.



Figure 6. Syd Field's three-act film structure: Act II is twice as long as Act I and III. The acts are separated by what Field defines as plot points. From *Screenplay: The foundations of screenwriting* (1994).

In analyzing the new Hollywood film (1960 or later), (Thompson, 1999) argues for a four-act film structure for an approximately two-hour long film (with five acts if the film is longer) separated by turning points instead of Field's plot points. Thompson's film segmentation into acts results in four thirty-minute segments. The difference in the two views comes mainly from the focus of the scholars' interests. Field's book is a screenplay manual and thus provides guidance on plot building to the plot builder. Thompson analyzes the structural make-up of the outcome of the plot builder's endeavor, the resulting feature film, with a focus on the characters' goals and change of goals as they are perceived by the viewer. The turning points tend to mark a change in characters' goal-oriented behavior. The first act is the exposition act, the bit of film that presents the narrative world in its current state. The second act is the stage for conflict and antagonism disrupting the exposition narrative world. Balance is regained through resolution afforded by the third act. In investigating viewers' trailer segmentation, the three-act structure seemed more appropriate due to the short duration of trailers. The

trailers in the corpus range from (“*Harvey* (film),” 2015) to the 3.44 minute *The Philadelphia Story* (1940). In the experimental procedure (described below), the acts are labeled Exposition, Complication, and Resolution.

Methods

Participants

Eighty-four participants segmented a total of twenty-one trailers using a computer interface. Each participant segmented three trailers with each segmented trailer belonging to a different decade and genre. The participants were provided with a user interface system that elicited a three-act structure of each film trailer they viewed. The duration and nature of each segment (as determined by each participant) were recorded. Parsers identified an average of 8.92 segments across trailers. One participant’s parsing of *Home alone* (1990) and *Valentine’s Day* (2010) were excluded for having 39 and 44 segments respectively (the mean being 8.92 segments with a standard deviation of 2.05).

Stimuli & Apparatus

The film trailer corpus used in this study consists of three trailers (roughly equivalent to the genres of drama, comedy, and action film) per decade from 1940 to 2010. With the decade of 1960 being represented only by the heist film *Ocean’s Eleven*, and the 1970 having only two films, *MASH* (comedy) and *Beneath the Planet of the Apes* (science fiction), the corpus contains twenty-one film trailers. The twenty-one trailers used as stimuli in this study are the IMDb official trailers of the films with the same title.

A user-interface program using the Apple Media Player application was constructed for this segmentation study. The user-interface displayed the film trailer on an Apple notebook computer screen. A grey bar gradually filled in the lower part of the screen and advanced as the trailer was ‘unfolding’ on the screen. The interface came with buttons the participant could click on to pause the trailer, to go back to the last segmented bit (a segmentation once done could not be undone), and to choose the type of segment the trailer displayed (the figure below for an illustration of the experimental setting). The segment type buttons (exposition, complication, resolution) could be pressed in any order and as many times as desired.

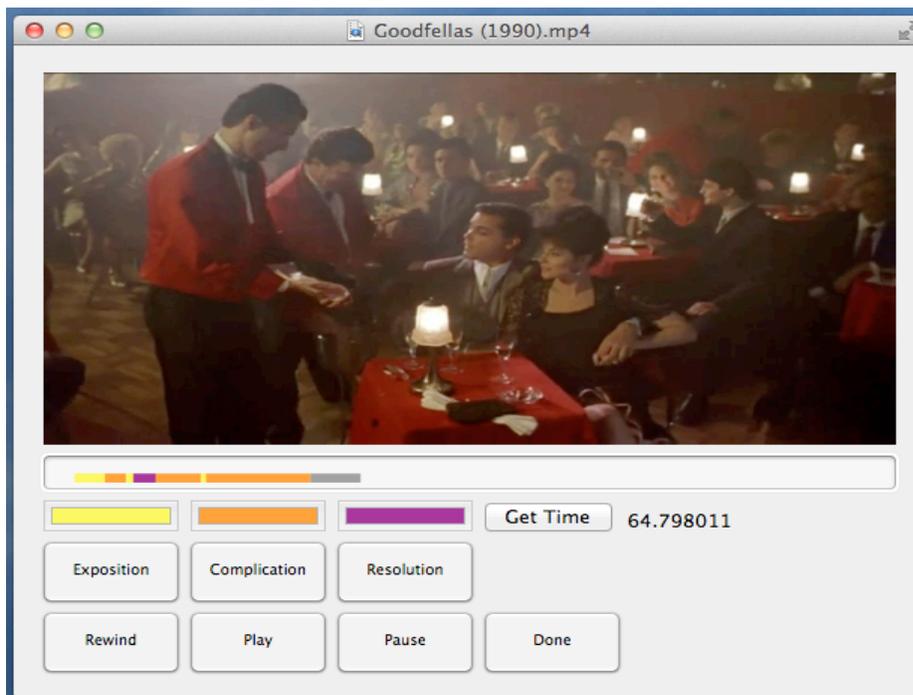


Figure 7. The trailer parsing experimental interface showing the grey bar as the trailer unfolds on the screen and the color-coded buttons for segment identification.

Participants were instructed to parse out the Exposition, Complication, and Resolution segments of the trailers with no constraints regarding the order of the segments or the number of times they press the same segment type button. The Exposition, Complication, and Resolution were defined in the instructions given to the participants as follows: the bits of film introducing the setting of the narrative with location and main characters (Exposition); the bits of film that indicate and/or undergo tension or conflict in the narrative world presented in the exposition (Complication), and the bits of film that announce or display events leading to a new balance in the initial world, the resolving of the conflict and re-setting of new norms (Resolution). The program output two files for each participant: a jpg file showing the sequence of segment types (see Figure 8 below) and a txt file indicating the number of strings each participant segmented the trailer into (i.e., every segment type button press was recorded as a number of strings with the strings being numbered and labeled in the text output file).

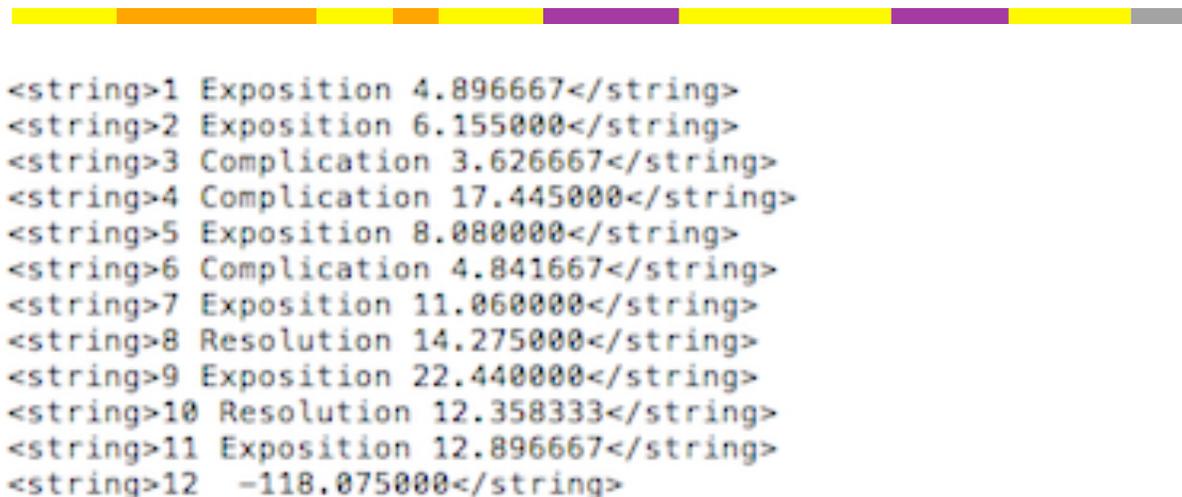
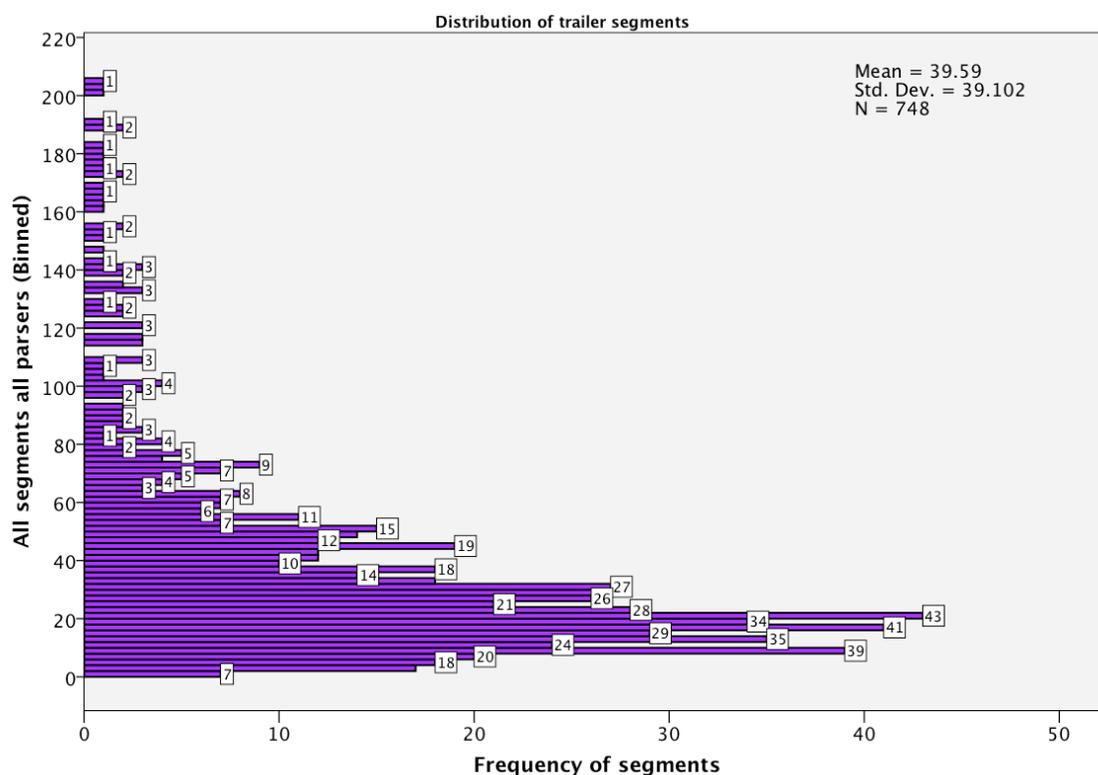


Figure 8. Above: The jpg output of the trailer parsing interface showing the identified segment types and their sequential order. Within each same-color bar there are multiple segmentations of the respective segment type. Below: The same-type strings are provided by the text file output.

Analyses and Results

The data were normalized using a unit-based normalization with the minimum value equal to the minimum duration of all segments (zero seconds), and the maximum value being the length of the longest trailer (*The Philadelphia Story*, 1940, 205.92 seconds).

Two histograms of the binned normalized data is shown below (Figures 8 & 9). Trailers are short sequences of film snippets. With the average trailer length in the current sample being two minutes and forty-two seconds (and the longest trailer being three minutes and forty-three seconds), the twentieth bin corresponds approximately to minute twenty-five in a film of one hundred minutes (the average length of feature film after 1940).



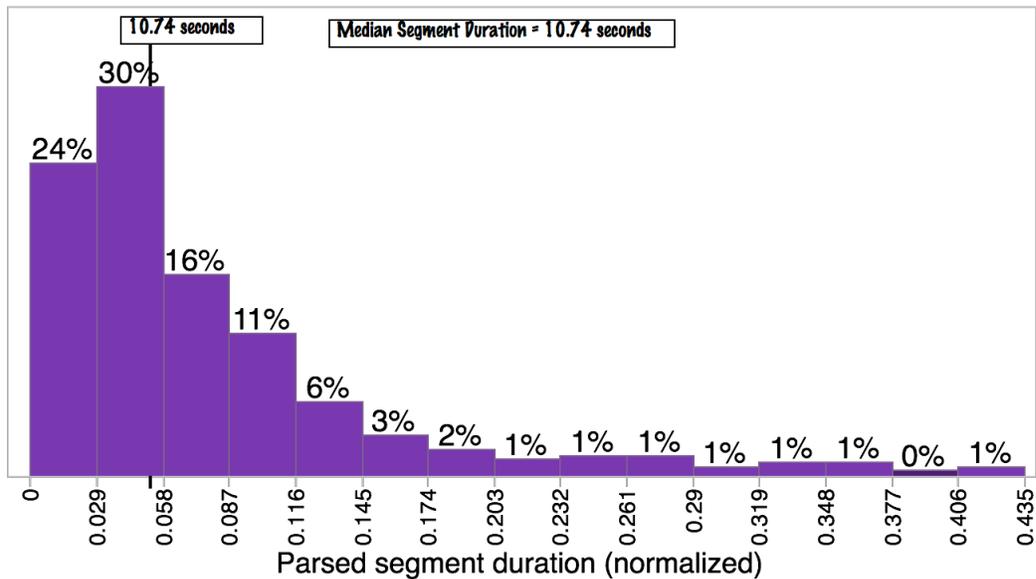


Figure 8 (above) & 9 (below). Positively skewed frequency distribution of segments per one-second bins: Parsers identified the highest number of segments (30% of total segments) in the beginning of the trailer (around the 20th second into the trailer). The segmentation becomes sparse as the trailer goes beyond the first minute (roughly half of trailer length on average).

Following the four-act film structure (Thompson, 1999), the minute 25 in a feature film of one hundred minutes would be at the end of the Setup act and at the beginning of the Complicating Action. That viewers' segmentations become finer-grained at the onset of the Complicating Action is most intuitive and reasonable: there are simply more events than in the setup. The shot durations in a full feature film show distributions that are conceivably consistent with the "rhythm" of the narrative as each genre has it. Action, for instance, may start slow with longer shot duration that decrease as the action picks up whereas drama may lengthen the shot duration during the complicating action that introduces tragedy and change. Cutting, Brunick and DeLong (2011) found that different

patterns of shot duration over the course of a film are a function of genre (see Figure below, from (J. E. Cutting et al., 2011)).

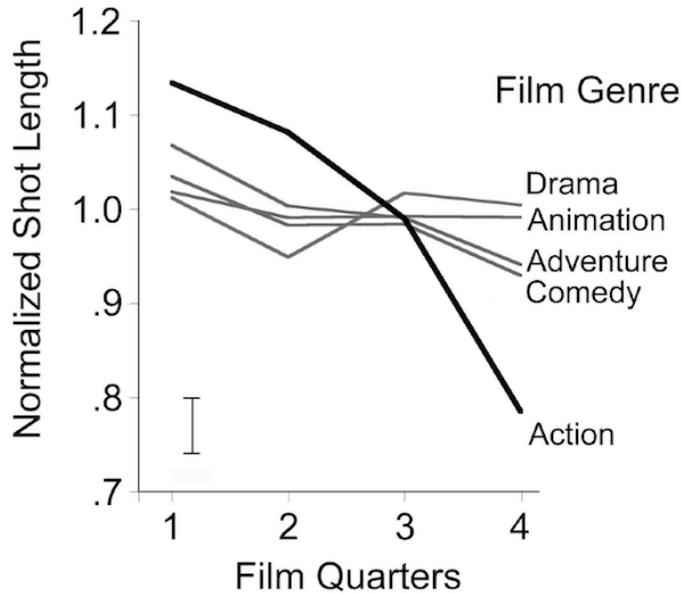


Figure 10. Shot length distribution across film quarters. From Cutting et al., 2011.

Films seem to have a fluctuating pattern of shot length distribution over the four film acts with the exception of the action film which shows a decreasing trend of shot duration from the film's beginning to its end. Despite the challenge posed by the variability across films due to different total length and number of shots (addressed in (J. E. Cutting et al., 2011)), the action film decreasing pattern will become relevant to the structure of action film trailers later on in this analysis.

Mission Impossible II, the Action film for the 2000 decade in the sample used in this study, shows a pattern of segmentation roughly equal to the structure of Action film over the course of the four acts illustrated in Cutting et al. (2011) above (Figure). *Inception* (2010), however, does not display the same descending pattern.

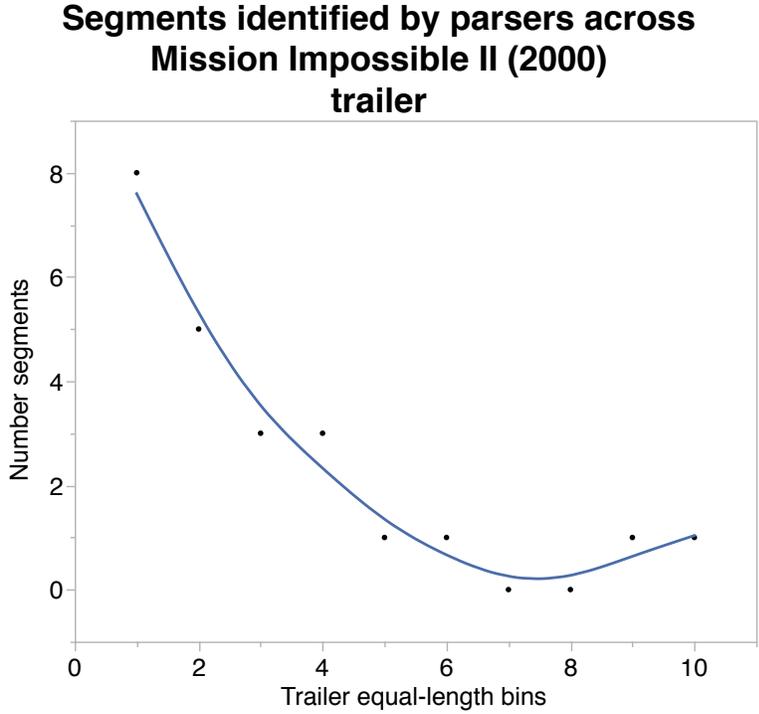


Figure 11. Parsers of the Action trailer *Mission Impossible II* (2000) identify more segments in the beginning of the trailer with a descending coarser-grain segmentation as the trailer advances.

Fleiss kappa interobserver agreement value of $k=0.43$ indicates a moderate agreement (Fleiss, 1971; Fleiss, Levin, & Paik, 2013). The average pairwise percent agreement of 74.09% suggests a fairly high agreement between any two parsers. The interobserver agreement and the pairwise percent agreement were measured on data binned in one-

second bins. In other words, 74.09% of every second of the film trailer two (out of four) parsers marked a segmentation boundary.

The viewers agreement on the segmentation of the full-length feature films based upon which the present trailer sample was constructed was 91% with an agreement value $\kappa = .56$ across 72 pairs of viewers (Cutting & Iricinschi, 2015). Both the agreement value and the pairwise percent agreement are lower in the case of trailers. Lower agreement on segmentation may speak to the complex, non-linear (without causally strong connections between adjacent events) nature of the stimulus.

But strangely enough, the viewers' segmentation of trailers across the seven decades does mimic the patterns Cutting et al. (2011) found for film of different genre. There is a sense in which the evolution of film over decades may be replicated within one film structurally. But a more plausible explanation (albeit somewhat related) has to do with the nature of the actual stimulus. We already know that the shot duration in recent film (approximately last 4 decades) has decreased leading to more frequent edits. In segmenting the drama film trailers in the current sample, viewers increase the rate of segmentation of the 2000 *Erin Brockovich* (Figure ...) and thus display a finer-grain segmentation with smaller segment lengths. The length of the segments identified for the 2010 *Inception* (Figure...), however, increases resulting in a coarser-grain segmentation. *Inception* (2010) should not be taken as a norm for any aspect of film making or film perception. A hard-to-follow plot can result in considerable variation in viewers segmentation for as long as we claim that the identified segments are "meaningful events".

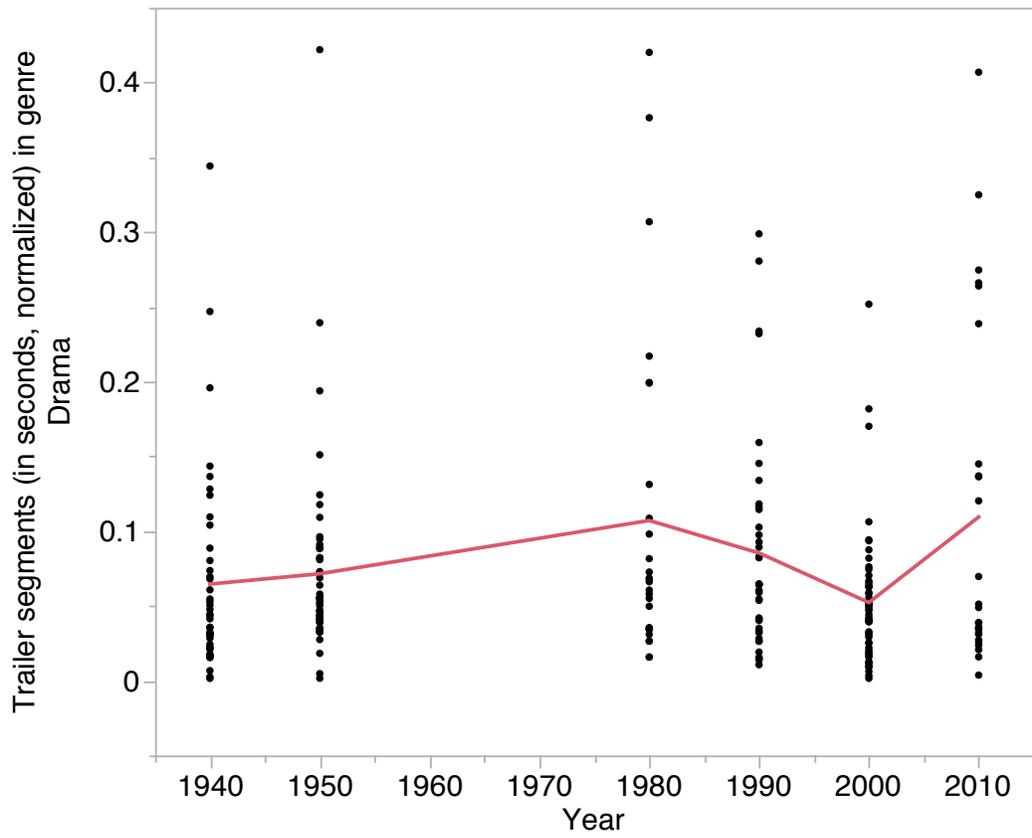


Figure 12. Viewers' segmentation of film trailers for drama films between 1940 and 2010.

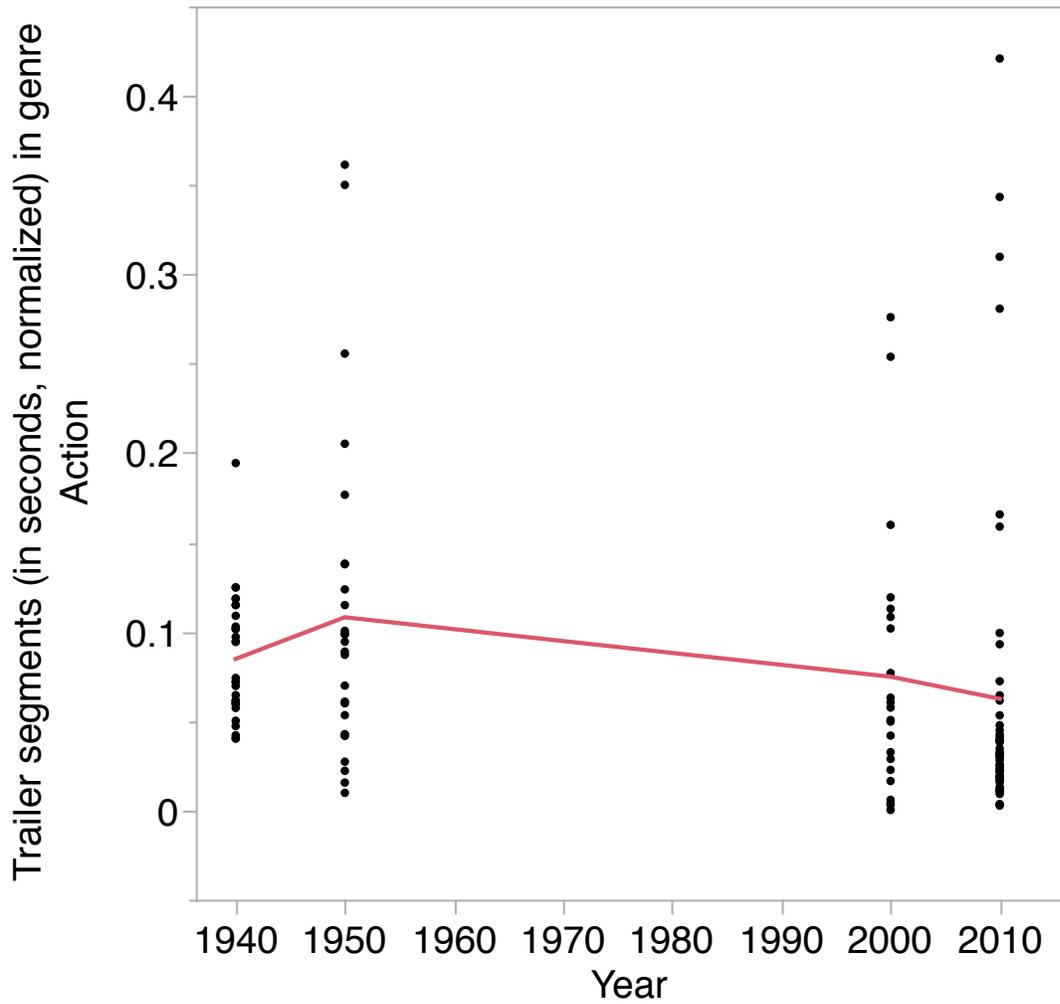


Figure 13. Viewers' segmentation of film trailers for Action films between 1940 and 2010. The length of the identified segments shows a decrease with the increase of the film release year.

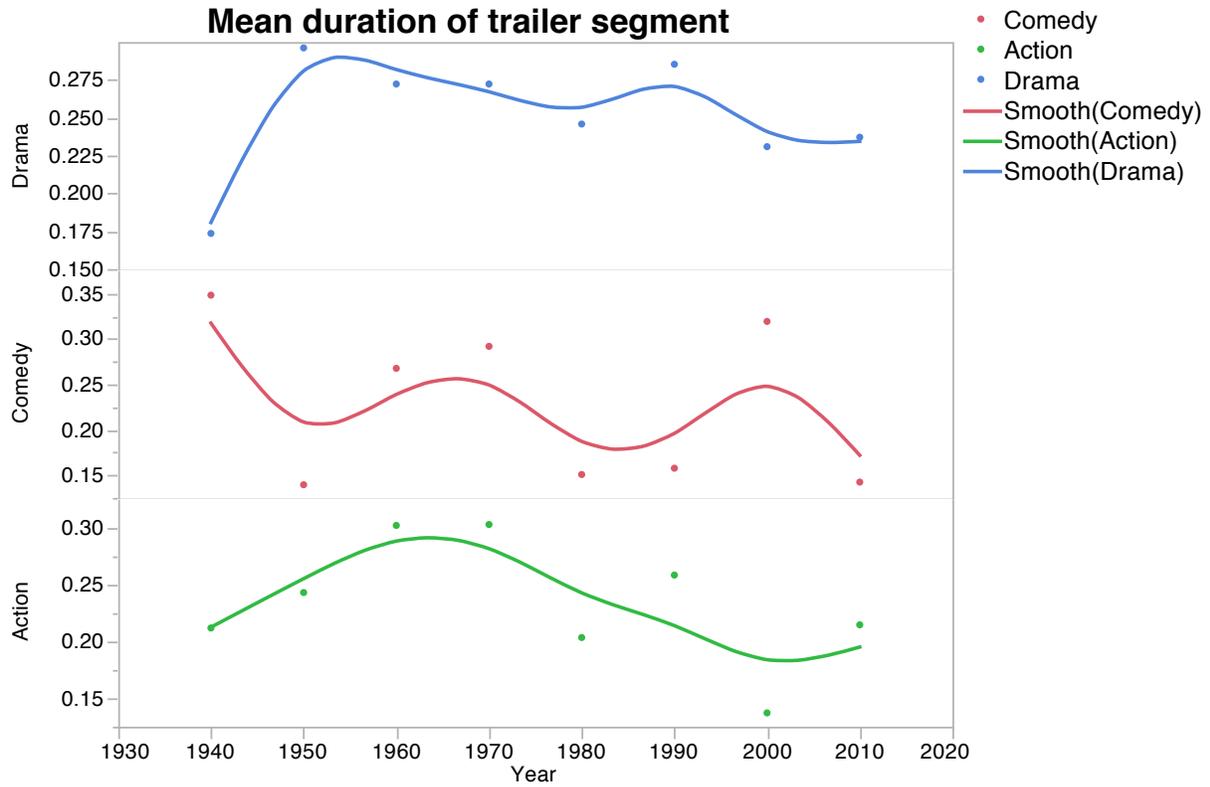


Figure 14. Number of identified segments in Action, Drama, and Comedy film trailers from the 1930-2010 decades. Action and Drama approximate a pattern of increase in segment duration followed by a decrease in the last two decades. Comedy does not display a reliable pattern.

It is tempting (and quite conceivable) to propose a “speed of thought” argument here (Pronin, Jacobs, & Wegner, 2008). Pronin et al. (2008) used various stimuli to “accelerate thought” and measure effects of the acceleration of mental activity. Fast-forward films were among the stimuli and task Pronin et al. used to make the mind race. Fast-forward films are artificially accelerated showing natural actions unfold at unnatural speeds. Film trailers may be thought of fast-forward films because of the high speed at which actual events are shown as well as the sudden shifts between unrelated events. It is therefore, reasonable to believe that the faster the stimulus (whether in low-level structural features or higher-level narrative content) the finer-grain the segmentation.

True to their promotional intent, film trailers entice viewers by withholding relevant narrative bits. The trailer introduces the viewer to the narrative world by presenting the setting and the characters depicted in the exposition. Stories revolve around change: Complication disrupts the narrative world, which, in turn, regains its balance after coming to terms with a major change. The resolution reveals one of the many possible futures the complication affords. Withholding most of the resolution creates a cliffhanger effect that will lure the viewer into the movie theater. The resolution segments are indeed the shortest (on average) in duration across parsers and trailers. The figures below show the Resolution as the least frequently occurring segment type in plots of segment duration as a function of segment type (see figures below).

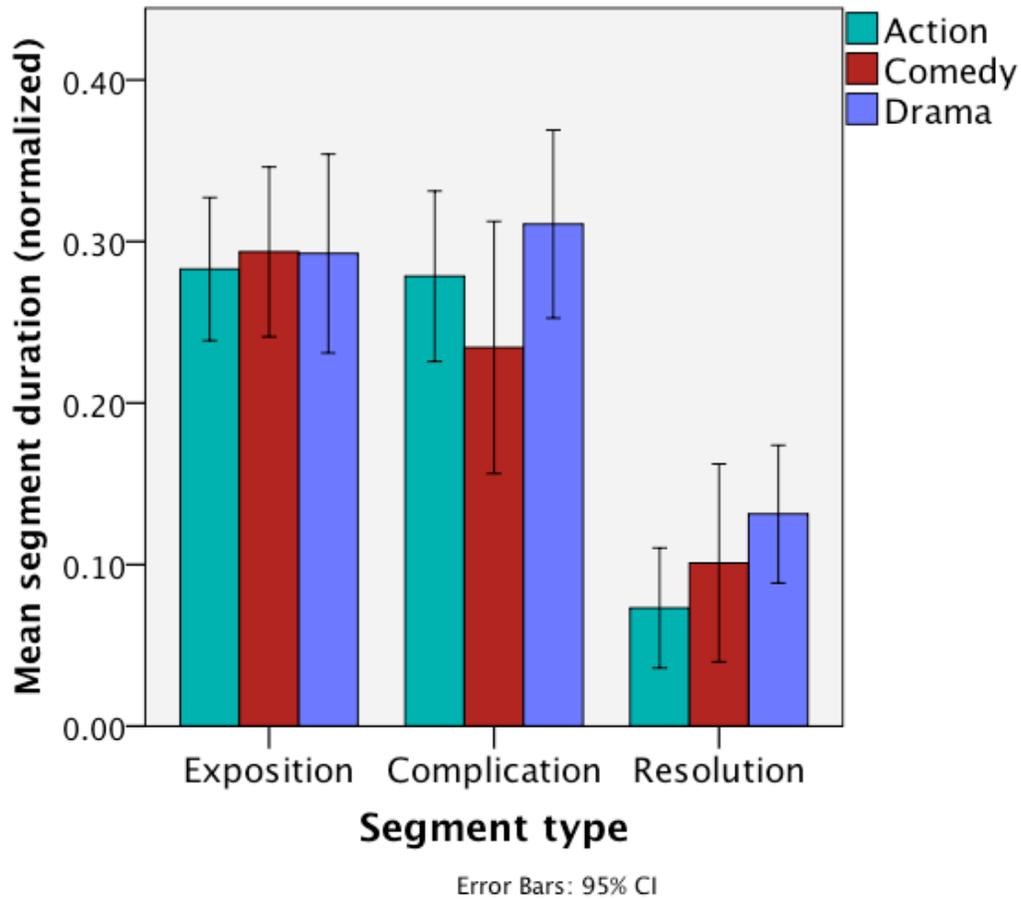


Figure 15. Mean duration for identified Exposition, Complication, and Resolution segments per trailer genre. The Resolution segment is almost absent in trailer of all three genres.

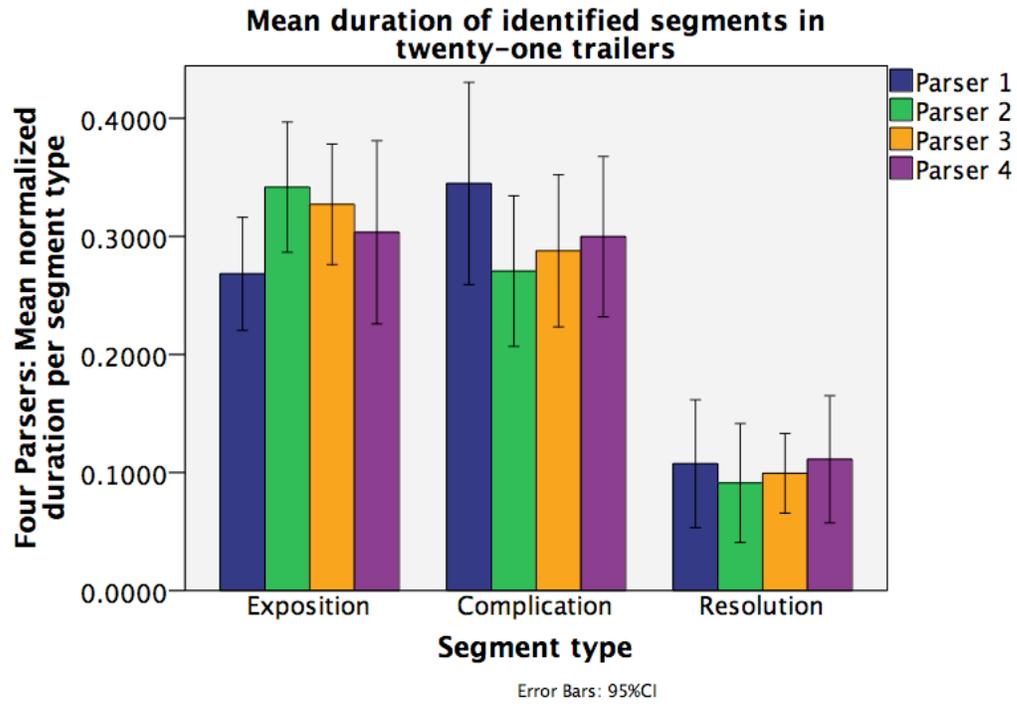


Figure 16. Duration of identified segments across parsers. All parsers agreed in finding a low incidence of Resolution units in trailers.

In conclusion, results indicate that, when asked to parse narrative labeled segments in film trailers, participants attribute most of the trailer time to Exposition and Complication, with the Exposition being on average higher than the Complication. The segment of Resolution is assigned less than 10% of trailer time. With the exception of Comedy, viewers of Drama and Action movie trailers segmented out more than 50% trailer time as the complication component. Older movies, however, tend to have longer Expositions with *Harvey* (1950), a true commercial dedicated only to describing and promoting the movie, are segmented entirely as Exposition.

The interobserver agreement is lower than the kappa value found by Cutting & Iricinschi (2015) for the segmentation of the full films promoted by the trailers in the current sample which indicates more variability between participants most likely due to the lack of causal connections between adjacent shots in the film trailer. *Ordinary People* (1980), however, presents the viewer with a highly structured 3-act schema in its trailer. It opens with the presentation of the narrative world, continues with bits of dramatic events, and ends with a false alarm for a resolution (see a more detailed discussion below). Participants segment the Exposition chunks in the beginning of the trailer, the Complication for almost the entire remainder of the trailer, and Resolution only after being primed by the main characters hugging towards the end of the trailer (second 149 out of 157). The graph below shows the density of segments the four parsers identified.

Segmentation of *Ordinary People* (drama, 1980)

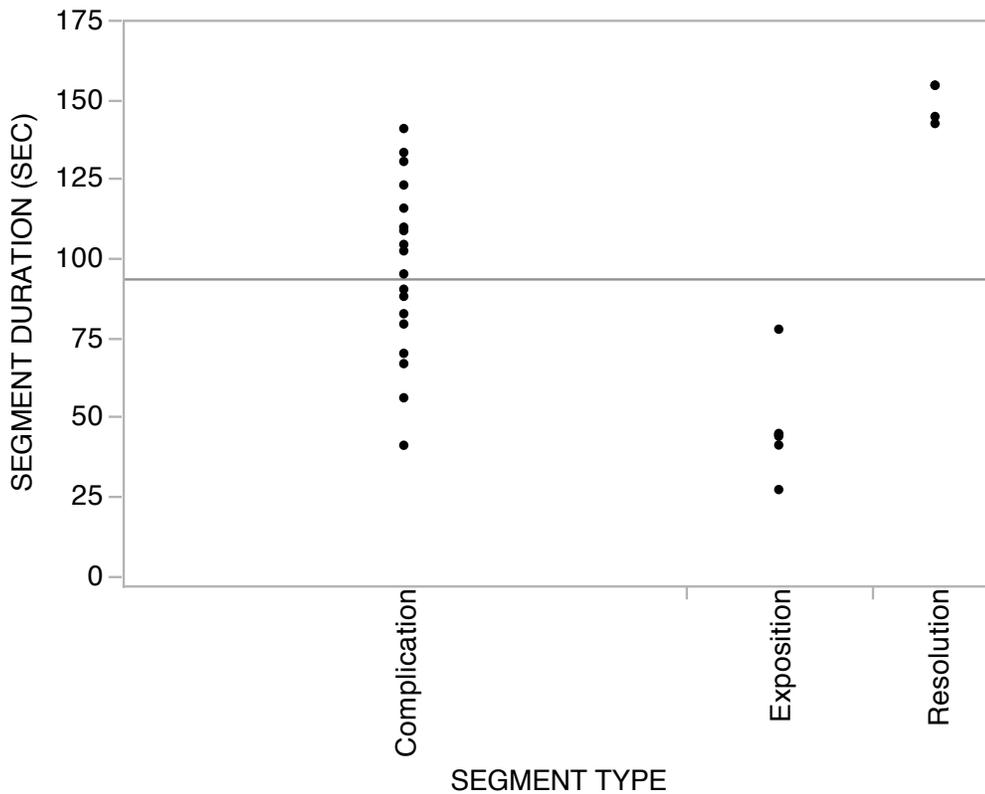


Figure 17. Segmentation of *Ordinary People* (1980). The Y axis represents segment duration as segments are parsed sequentially along the duration of the trailer. Total trailer length is 156 seconds. The Exposition, Complication, and Resolution segments are identified by parsers in the sequential order of a three-act structure.

Discussion

Trailer viewers agree on their segmentation 79% of the time in pairwise comparisons.

The interobserver agreement kappa value indicates a moderate agreement across all parsers. We can therefore conclude that despite the discontinuous nature of the trailer, viewers can still identify narrative bits. There are many methodological aspects that need a closer look and analysis. First, the task is an approximation of a forced-choice task in that the nature of the segments is pre-established by having the labels provided for the segmentation. This may explain the fairly high level of agreement. Data on trailer

segmentation without the provided labels would clarify this issue. Secondly, the shorter shot duration in trailers compared to full-length film may elicit a finer-grain segmentation. Existing research (Swallow et al., 2010b; Zacks, Speer, & Reynolds, 2009) asked participants to segment with the smaller unit of action (or event) in mind as well as with the largest unit of action in mind. The two sets of instructions resulted in a fine-grain and a coarse-grain segmentation. Trailer and full-length film may be thought of as eliciting fine-grain and coarse-grain segmentation respectively simply because the shots are consistent with a finer or coarser parsing. And finally, trailers are commercials indeed and in their promotional intent they “communicate [their] film and at the same time it [have] to hold it back” (Jensen, 2014). The data presented above suggest that the most efficient strategy for a trailer to reveal just the right amount and type of diegetic context is to present the narrative setting and complication tension while hiding the resolution that normally contains the valence of the ending thus indicating whether the movie is a happy-ending story or not. Films based on true stories, such as *Erin Brockovich* (2000), take a lesser risk in presenting the happy ending of the already-known story whereas fiction dramas (such as *Ordinary People*, 1980) hone in on the emotion of and tension among the characters without revealing the nature of the resolution. If anything, dramas may give the viewer erroneous cues to misdirect the anticipation of the ending. The illustration below presents a sequence of frame extracted from the trailer for the *Ordinary People* (1980). The sequence maintained the order in which the scenes corresponding to the frames below are shown in the trailer.

An illustration of the “intensified emotion” (Jensen, 2014) that trailers of drama tend to elicit is provided below as a sequence of frame from *Ordinary People* (1980).

The illustration presents a sequence of frame extracted from the trailer for *Ordinary People* (1980). The sequence maintained the order in which the scenes corresponding to the frames below are shown in the trailer.

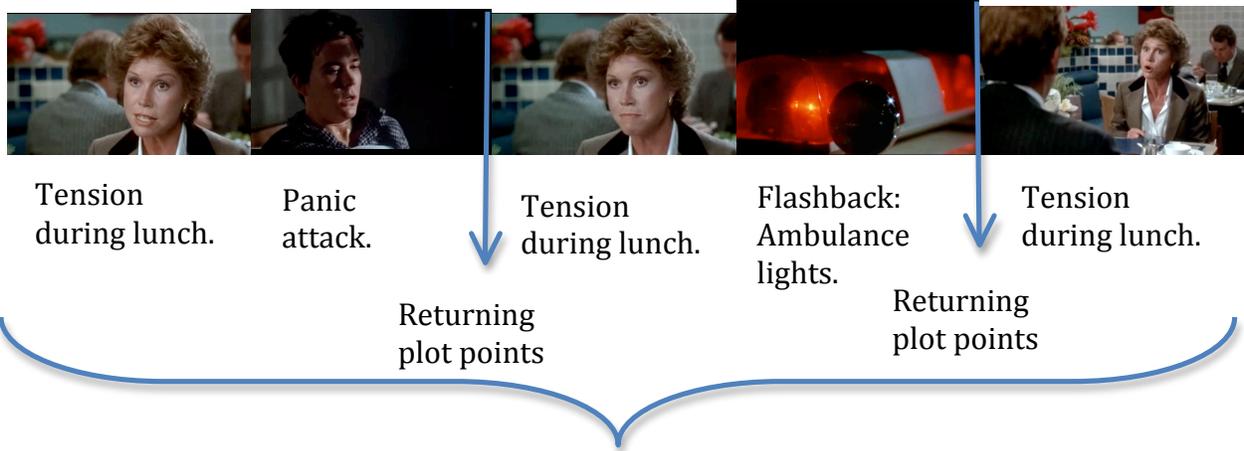


Duration: 15 seconds

Establishing shots in the beginning of the trailer introducing the narrative setting. This sequence is followed by a black screen fades in to credits. It is the Exposition segment.



Tension introduced immediately after the credits. Beginning of the Complication segment.



The Complication segment. The sequence above contains six shots, each approximately two seconds long.



Figure 18. Continuation of the sequence above: Flashback to the son's suicidal attempt.



Figure 19. Last frame in the trailer is a Resolution false alarm. The scene misguides the viewer into inferring a concluding false resolution in the film.

Commencing with classic establishing shots that present the narrative setting and the main characters, the trailer for *Ordinary people* switches abruptly to tension and tragedy shown in a rapid succession of brief shots. In *Ordinary People* (1980) the incidence of Complication segments identified by participants is considerably higher than the incidence of the other two segments. The establishing shots in the beginning of the trailer take the first 50 seconds of the total trailer time. Three of the four Exposition

segments are identified within this introductory beginning. Consistent with the genre of drama, 64.2% of all segments participants identified are under the Complication label. As for the Resolution, participants identify Resolution segments only after the scene (around second 150) showing the husband and wife hugging (see figure 17 below). Although the scene exists in the full-length film, such reconciliation does not occur in the film. The trailer creates a new narrative by ending on a resolution note.

Not only does *Ordinary People* use the black screens that trailers generally use to separate chunks of narrative, but it also interweaves bits of obviously coherent events. One of the sequences I present in Figure 18 above has two returning plot points in two seconds of trailer. The tense lunch scene that features the couple discussing tragic family issues while attempting to have lunch at the mall alternates rapidly with flashbacks referring to the son's suicidal attempt. The lunch scene thus becomes three scenes over the span of a couple seconds.

In reference to continuity editing, Bordwell and Thompson (2004, p. 310) contend that “the purpose of the continuity system is to create a smooth flow from shot to shot”. When it comes to trailers, the principles of continuity are deliberately violated to create a narrative with gaps that only the viewing of the full film can fill. Film edits consistent with continuity rules result in high agreement among parsers. Trailer parsers agree only on the absence of the resolution segment from the core-narrative constructed by trailers. Analyses on parsing unlabeled trailer segments would clarify the nature of agreement or disagreement among parsers.

1.3.2. Parsing puzzle film: Narrative content or perceptual salience?

In the section above, I claimed that trailers are complex stimuli; at least more complex than the full-length feature film they try to sell. The complexity in the case of the trailers may have more to do with the nature of the task as it interacts with the trailer structural make-up. What I mean by this is the following: Trailers moves fast and they jitter a great deal. They jitter with black screens inserted in the middle of an intense scene, they jitter with narrator voice calmly introducing the film while in the background Erin Brockovich is experiencing a major life crisis, they jitter with a multidimensional domain in which the dimensions do not align. The task of segmenting such a stimulus is bound to have some extraneous noise. Some of this jitter has become recently part of full-length feature films. Whereas older film obeys rules of continuity editing, younger film tends to break such rules deliberately. “Continuity editing, the dominant form of an interlocking set of classical editing techniques is created through handling space and time, [...], binding shots within units, [...]. What happens *across shots* within a unit often occurs within more or less continuous time and contiguous space; what happens *across units* often occurs across initially unknown *gaps of time and space*” (J. E. Cutting, Brunick, & Candan, 2012). The gaps of time and space can be implemented in the structure of film in ways that render the narration easier or harder to comprehend. The complex structure of film can be induced by reversed temporal orders that break the causality links between events (as Christopher Nolan does in *Memento*, 2001) or introduce multiple temporal and spatial orders with intricate dynamics among them (as Christopher Nolan does in *Inception*, 2010). Given that our cognitive processes are

anchored in space and time contingencies, such films make for complex stimuli worth looking into.

This section investigates the impact of the complexity of a film structure on the way viewers parse cinematic narratives into distinct events. I will first discuss the structural aspects that create a complex narrative by establishing analogies with challenging language processing and spatial cognition tasks. *Memento* and *Inception* (Nolan, 2000, 2010) are relevant case studies due to the complex structures they introduce. Parsing data suggest that viewers rely on perceptual salience in segmenting complex film narratives.

Parsing perceptually continuous narratives into distinct events is a prerequisite for reconstructing the *fabula* – the re-organization of narrative events into a temporally, spatially, and causally coherent sequence – from the *syuzhet* – the actual display of the narrative that departs from the natural chronological, spatial and causal consistency. Empirical evidence reveals a high degree of consistency across viewers in parsing film narrative (Cutting, 2012; Magliano & Zacks, 2011). These studies, however, tend to use linear narratives as stimuli: narratives that make the temporal and causal relationships accessible to the viewer during one single viewing endeavor. Reconstructing the inconspicuous *fabula* from the visible *syuzhet* requires “assumptions and inferences” (Bordwell, 1985) that are challenging in the case of non-linear plots. How do viewers parse a film into events when such assumptions and inferences do not yield an immediate coherent pattern? Narratives seem to continuously increase in complexity and thus demand more of the viewers’ cognitive engagement. Puzzle plots weaken the

connection between the fabula and the syuzhet thus making the fabula less inferable during the traditional single movie-theater viewing. Two relatively recent films, *Memento* and *Inception* (Nolan, 2000, 2010) ‘scramble’ – albeit following coherent principles – the temporal and causal relationships between events. The scrambling of the causal order of events results in puzzling plots that make the fabula (and thus the narrative comprehension) almost inaccessible to the viewer.

What makes puzzle films puzzling? I will elaborate on three processes to account for the processing complexity in puzzle plots: Center embedding, interference, and mental rotation.

A theory held by certain areas of linguistics asserts that recursive language structures are not constrained syntactically. The sentence *The rat the cat the dog chased killed ate the malt* is thus perfectly acceptable on theoretical grounds. The sentence has three layers of center-embedded clauses: *the dog chased the cat*, *the cat killed the rat*, but not before *the rat ate the malt*. The speakers’ perspective, however, departs from the perfectly acceptable view and acknowledges the constraints imposed by actual language use. Language speakers find sentences with three levels of embedding hard to comprehend, cognitively taxing, and highly contingent upon expertise (Christiansen & MacDonald, 2009). Recursive levels of embedding have been used in narrative plots with the same apparent demands on online processing.

With its complex structure, *Inception* presents the viewer with three levels of embedded dreams strangely interconnected. The cognitive challenge of the viewer consists of maintaining each level in her/his memory throughout the film and connecting the levels

into a coherent puzzle. Moreover, the three levels of dreams (with limbo and reality as additional alternatives) follow different time orders and are presented in alternated sequences. Similar to the centrally embedded sentences, viewers find dis-embedding the three dream levels challenging when watching *Inception* (see Figure 20 for illustration).

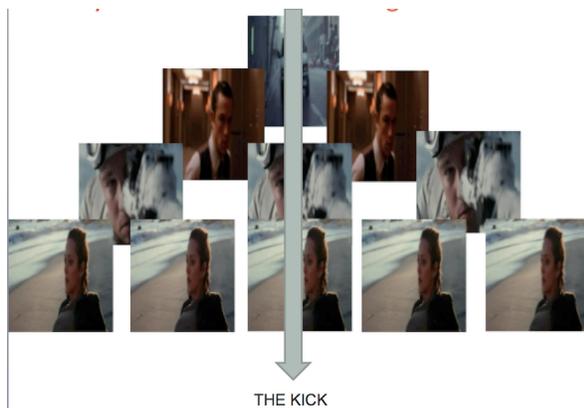


Figure 20. The center-embedded structure of *Inception* (2010).

Four participants watched the film in its entirety and identified events based on their own criteria. To avoid biases, no constraints were provided to the participants prior to their parsing task. Only 10.48% of all identified event boundaries showed agreement between two (any two) of the four participants while 7.48% were one frame apart. Participants were also asked to provide brief descriptions of the events identified. Judging by these descriptions, the viewers of *Inception* failed to connect the events in a meaningful way; “Now sitting in office”, “man sitting in a restaurant”, “man driving while being shot at” are a sample of the approximations the parsers provided while segmenting the film. Three elements define the parsing consistency in the case of *Inception*: the soundtrack (diegetic and non-diegetic music increasing in amplitude), the onset and offset of the inception operation on the airplane, and visually conspicuous

spatial changes (i.e., from room to white mountain). Parsers used perceptually salient features (e.g., soundtrack, color) instead of content-related markers as event boundaries. *Memento* presents the viewer with two time orders that unfold in opposite directions (see Figure 21).

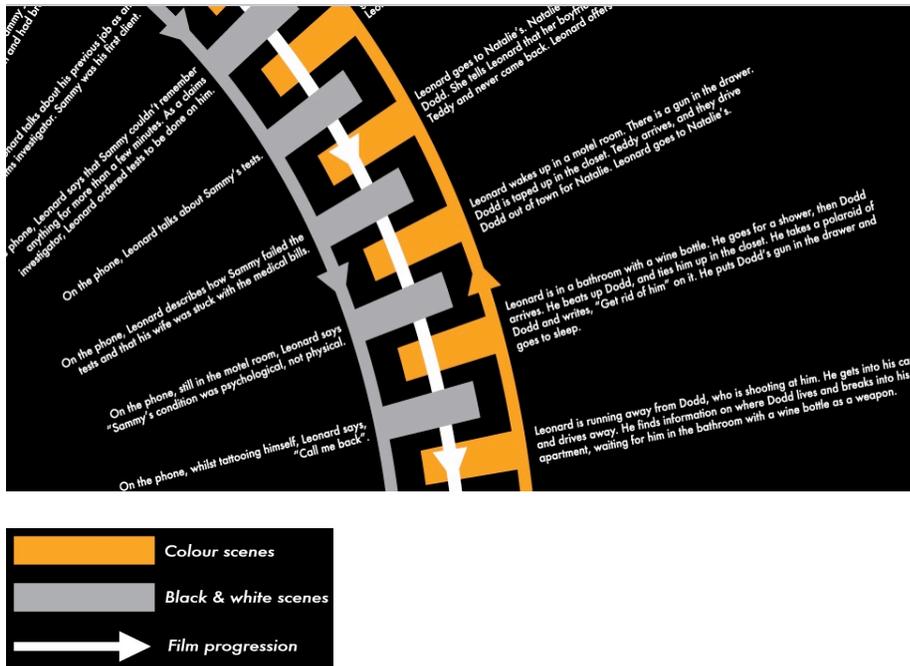


Figure 21. Diagram of two opposite time orders in *Memento* (2000).

The film consists of 22 black-and-white segments - narrating events in chronological order - alternating with color sequences that narrate the events in reversed temporal order. Moreover, every color sequence ends with a shot that opens the following color sequence. The reversed temporal order affects causality: earlier events are effects and not causes of subsequent events. *Memento* challenges the memory of the viewer by alternating two visual tasks, the color and the black-and-white sequences. Consistent with research on interference and memory, *Memento* makes it difficult for the viewers to

form and update situational models for the narrated events. In order to arrive at the fabula, the viewer has to record the matching shots that mark the color sequences and “perform a continuous mental rotation by putting the first sequence into the final position” (Ghislotti, 2009).

Memento has a clear formula that encodes its structure and makes it appear rather algorithmic, devoid of random, functionless artistic bits. The formula simply encodes the well-orchestrated sequence of black and white and color shots: C1.ms1 + BW1 + ms2.C2.ms1 + BW2 + ms3.C3.ms2 +...Color shots are followed by matching shots (i.e., identical shots that mark the two time order) that in turn are followed by black and white shots. Color shots advance in natural order while the black and white shots go backward in time. The challenge does not consist only in disentangling the interleaved color-black&white sequence, but also in the mental rotation required by color sequences. A sequence such as ms2.C2.ms1, once rotated would provide a continuity cue by reiterating the matching shot (see Figure 22 below).



Figure 22. Matching color shots in *Memento* (2000) signal (if noticed) the reverse temporal orders.

There are 22 sequences of each kind in the film – see Figure 23.

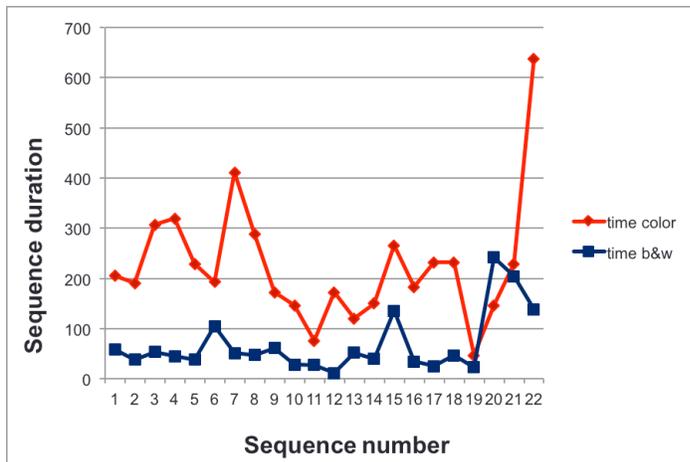


Figure 23. Duration in seconds of the 22 black & white and 22 color sequences in *Memento*. The color sequences are consistently longer and carry more of the narrative content.

Four participants were instructed to segment *Memento* into events. The transitions between black-&-white and color are marked by black screens preceded by fades at the onset and offset of black-and-white sequences. All such transitions were marked as event boundaries by all parsers. Participants were asked to provide a summary of the film after parsing it. The summaries were accurate and coherent, with only two participants (who were familiar with the film) mentioning the two time orders the narrative contains.

Despite the obvious comprehension difficulty, all participants reported having enjoyed the challenge of the puzzle films even at the expense of comprehension. The modern viewer seems to prefer the more engaging riddle-like film narrative that requires several viewings. Puzzle films part with the passive moviegoer and engage the video gamer's persistent habits.

2. Dimensions of Events: Space, Culture, and Stereotypes in Multi-Cultural Film

The studies I have presented so far dealt with book readers' and film viewers' sensitivity to spatial and temporal changes in narrative. Space and time, where narrated or directly experienced, mark all our experiential life. Space perception deserves a closer look. The abstract, elusive time is, in the way we encode it and cognize about it, subordinate to the concrete visually and tactilely accessible space: we use a spatial framework to make time cognitively manageable.

Buckland (2009) opens his edited book on puzzle plots with a note on culture.

“People from all cultures understand their experiences and identities by engaging the stories of others, and by constructing their own stories. But in today’s culture dominated by new media, experiences are becoming increasingly ambiguous and fragmented” (Buckland, 2009, p. 1). The new media Buckland refers to continuously receives tremendous scientific attention. When discussing space, it is most likely the virtual space we cyber-fight over. When discussing violence, it is most likely video gaming that that is the culprit.

In the section below, I argue that space, the good old space as we know it, the physical tangible space, can affect us in subtle ways when presented on the screen. Film is a cultural artifact and I study it as such in the section below. Space can be (or could be in the past) equated with culture since it used to mark cultural regions. I will analyze the case space in cross-cultural film by looking at films from three different cultures (US, Israeli, Arab) and the way they present their spaces and each other spaces. Segmentation data of a sample of these films will be briefly discussed.

A pervasive and heterogeneous presence, space has attracted various definitions in multiple domains of scholarship. For political scientists and philosophers, space may equate power; for cognitive scientists space provides blueprints for memory organization, and psycholinguists contend that spatial structures assist in conceptualizing the elusiveness of time. This essay discusses the semiotic of space and its visual presentations in cinematic narrative. Filmmakers situate their narrative content in space and time and thus use spatial representations as a dimension of narrative discourse.

Even when strictly confined to depictions in film narrative, space has received a rather intimidating variety of interpretations. Facing such complexity, the question, although simplistic and generic sounding, “what is space”, would guide and constrain the contextual search for the significances of space. Tuan (1977) answers the question – that he asks verbatim – by attributing an experiential basis to our knowledge of space. Seen as experience, space perception implies contextual dependency and a degree of peculiarity with each perceptual instance of each individual. The present essay alters the generic ‘what is space’ question and formulates a new - only slightly less generic – experiential concern: What is space in the experience of film viewing? With film being a visual artifact that reliably uses space to signify, decoding filmic spaces may appear as an intimidating semiotic morass. However, viewers’ experiences with interpreting spatial layouts on the big screen seem to be constrained by filming techniques (i.e. camera

angle) and cultural background. It is at the level of contextual similarity across cultures and high/low camera angle shots that filmic space is analyzed in this essay.

In what follows, I will briefly outline the relevant interpretations of space in film proposed by various film scholars. Space will then be discussed as a dimension within a culturally-connoted cinematic discourse. With culture as a contextual constraint, the analyses of spatial perception induced by film viewing reveal inter-cultural dynamics that the increasingly universal film industry affords.

2.1. Spaces in film: Filmic space, narrative space, and viewers' engagement

Understanding space as presented in film is as beneficial a skill as navigating space in real world: It enables us to encode the relevant boundaries and reference points on a cognitively manageable narrative map. Our environment displays an inordinate array of structures and patterns configuring spaces and places. Experientially and in terms of social behavior, places suggest safety and belonging, whereas spaces allude to freedom and escaping (Tuan, 1977, 2003).

What is space as represented (and experienced) in film? *How* is space represented (and experienced) in film? Do visual presentations of space play any role in viewers' engagement with the narrative? Do they affect the viewers' 'being there' sensation?

As a visual form of art, cinema simply shows places and spaces, it does not (in Chatman's terms, 1980) assert or describe them. In showing space, film makes spaces

directly available to the viewer's eye, thus creating the illusion of non-mediated perception. Depictions of space in film are an inherent part of the cinematic narrative, subject to the filmmaker's crafting to the same extent as any other framing or contextualization aspect. The film spectators, in turn, attribute connotations and significances to spatial layouts to the same extent they do to gestural behaviors of characters.

The most intuitive (albeit simplistic) definition of space in film makes the screen approximate a 'container' metaphor: Space is contained in the film frame or left out of the film frame (Burch, 1973). The projection screen thus becomes the confining physical element that determines visible filmic space. Burch identified six off-screen spaces: left, right, top, and bottom of the screen, beyond the set and behind the camera. Obviously, the viewer focuses on the on-screen spaces. The space excluded from the shown screen, however, suggests possible relationships and contingencies between filmmakers, film viewers, actors, and the narrative. In the vertical plane, the four sides of the screen 'hide' spaces that may be accurately inferred by the viewer as logical continuations of on-screen space. In the horizontal plane, the area behind the camera affords little-to-no narrative content, whereas the beyond-the-set area may allow for the development of storyline bits⁷. I am tempted to claim that the on-screen spatial enclosures together with their inferred extensions create the narrative space: Seen by both actors and film viewers, imagined, dreamed, or reminisced by characters, the narrative space binds all narrative loci and constructs a story world. The filmic space

⁷ "The use of the fifth zone of off-screen space, that behind the set, is, of course, common; characters go out a door and are now concealed by a wall or a staircase." (Bordwell & Thompson, 2004, p. 259). But even when out of sight these characters may contribute to narrative development.

and the narrative space engage the viewer in a more complex process. The filmic space highlights the relevant narrative loci (thus attracting viewer's gaze) and affects the audience's engagement with the narrative. When such filmic spaces are 'weak'⁸ and void of narrative content, the audience disengages. Moreover – as I will show in what follows – the filmic space provides information for the extraction of a more abstract narrative space, one that maps plot elements onto a dynamic virtual space, rather than actual locations.

2.2. Spatial Configurations and Narrative Maps

Spatial adjacencies and juxtapositions, as well as spatial separations and boundaries – indeed, all relationships established across spaces in general -- are narrative elements in their own right. Behaviors of characters that share a space are more conducive to interpretations consistent with contingency and joint narrative roles than are behaviors separated in space⁹. So, what is space in the film experience?

In discussing classical film, Bordwell, Staiger, and Thompson (1985) describe space in early Hollywood film as the background structure supporting the character-based narrative causality: “In making narrative causality the dominant system in the film's total form, the classical Hollywood cinema chooses to subordinate space” (Bordwell, Staiger, Thompson, 1985, 50). The *filmic space*, as the authors call it, is thus relevant only to composition and framing techniques

⁸ I will define and illustrate weak filmic spaces later in the essay.

⁹ As an illustration, Cutting, Iricinschi, and Brunick's (2013) research analyzing screen spatial contingencies is elaborated upon below.

used to 'populate' it. The focus on composition in filmic space approximates the 'space as container' metaphor proposed by language researchers: Norms regarding centering, balancing, and depth insinuate an initially empty frame that is 'filled' – just like a container – with narrative items. The space containees tend to be the characters that are placed centrally (according to the centering rules followed in early Hollywood productions), and are emphasized or de-emphasized by lighting. By appropriating the 'filmic space as container' metaphor, space becomes removable and re-usable, a detached area used to palimpsest narrative content over. Historically, composition-related filmic space became relevant when film scholars started to address the transition from single-shot to multi-shot film (Salt, 2009): The composition of the frame space within one shot afforded consistent point-of-view shifts in the following shot and maintained spatial coherence across same-scene shots. I would confidently argue that presentations of spatial layouts in film have consistently received less scholarly attention per se because of this initial focus on space as only relevant to editing techniques. As Thomas Elsaesser (1990) observes, the interest in the early cinematic endeavors was motivated primarily by the technological advances and the machinery utilized to enable multi-shot film, and not necessarily by the artistic dimension of the film form. Space was thus confined to the realm of filmmaking technicalities. Filmic space, however, can be (and indeed came to be) an inherent dimension of the cinematic narrative, a dimension with its own semiotic. As an inherent dimension, filmic space affords various mappings that delineate and chart out distinct narrative spaces.

Let us revisit the space as container metaphor and conception of the film frame as an empty space that welcomes its characters as they 'populate' the screen throughout the film. Cutting et al. (2012) uses screen-as-space information in a novel way that is not inherently linked to editing techniques. Using a computed distance between characters – where the smallest distance links the characters that share the highest amount of screen time – Cutting et al. created narrative maps that encode relationships between characters as they develop throughout the film. Imagine a plot based on a marriage being disrupted by an extramarital affair. Cutting et al's narrative map would initially plot two adjacent dots representing the two spouses before the affair. As the film unfolds, the map will show the development of the characters with the distance between the two spouses gradually increasing as the affair-inducing character attracts one of the spouses closer and closer. Well, now -- the seasoned movie-goer may say -- how is this not entirely intuitive? After all, behaviors of characters that share the space are more conducive to contingency interpretations than behaviors separated in space. The movie-goer, however, no matter how seasoned, would not intuit that the mere screen time encodes character development and character change. It is hard to believe that screen time shared with other characters – and nothing but shared screen time, since the sound track and subtitles were not coded – would provide viewers with enough information for them to infer reliably such character developments. But it does. Filmed in one single location with the characters placed in the same order around a table, *12 Angry Men* (1957) serves as a rather extreme example of such narrative maps. Even in this stationary setting, the shared screen time indicates the change of jurors' mind

regarding the guilty/not guilty verdict. The jurors who voted 'not guilty' cluster together with the distances between them between growing considerably smaller as the film progressed from the initial setting.

Given how sensitive we – as cognizing entities -- are to temporal and spatial contingencies (developmentally, e.g., Watson, 1966; and as learning mechanisms in general, e.g., Bassili, 1976, Meltzoff et al., 2009) shared screen time assists viewers in constructing expectations while processing the story unfolding on the screen. A vast body of research on statistical learning (e.g., Saffran et al. 1996, Conway & Christiansen, 2005) indicates that the human mind processes stimuli by keeping track of co-occurrences and thus developing probability-based expectations. Cutting et al.'s analyses indicate that film viewers have sufficient visual information to anticipate the film's character dynamic and map the character change on an abstract narrative map. This proposal dissociates fairly sharply the filmic space from the narrative space by advancing an abstract narrative domain. While filmic space here abides by Bordwell's initially-empty container metaphor, the narrative space simply encodes character change and between-character relationships based on the visual input alone. As proposed by Cutting et al., the narrative space is not a physical locus marked by clear boundaries, but an abstract domain that hosts the characters' dynamic as they create the narrative. The visual information provided by the filmic space sustains the abstract dynamic narrative space. The field of linguistics (certain domains of it) insists on distinguishing between structure and meaning, with meaning being inferred from the stand-alone structural information. The blue print for such meaning can be encoded as

‘who did what to whom’ – essentially, the relevant bits of communicative acts, agency and its outcome. As a potentially explanatory analogy, the blue print for Cutting et al.’s narrative space could be encoded as (the humorous connotation notwithstanding) ‘who did what *with* whom and for how long’. This paradigm affords extracting the story through statistical regularities (i.e., patterns of co-occurrences) established by character interaction. The structural aspects of filmic space, such as spatial and temporal orders (i.e., spatial and temporal contingencies) establish a dynamic relationship with the narrative space by affording pattern abstraction exclusively from characters’ shared screen time, as well as alteration and online update of such patterns. The two spaces – filmic and narrative – are thus in a constant informational exchange. It is precisely this dynamic informational exchange that engages the viewer and affords the ‘being there’ impression.

2.1.1. Patterns in narration: Views from above in visual media

As argued above, filmic space and narrative space engage each other in a dynamic informational exchange. And by engaging each other in this dynamic process, they engage the viewer in a problem-solving task of anticipating and updating storylines. What if this dynamic informational commerce between the filmic and narrative spaces is disrupted? What if the filmmaker minimizes the content of the filmic space (as a place holder) to such a degree that the narrative domain receives a far too underspecified input? What narrative space would the viewers’ construct then?

For as long as film claims to approximate a comprehensible visual language, images that replicate the natural eye-level approach should be the obvious trendsetter. And indeed they are. The alternative, views from above (with the bird's eye view as the extreme version) and views from below (with the worm's eye view as the extreme version), are deemed unnatural and difficult to process. The motion picture eases the viewers into the decoding process of such high angle camera shots by preceding them with relevant contextual information. "Marked departures from a horizontal lens axis often produces images which can be rendered comprehensible and acceptable in films because the activities in them are already understood from the previous movements of the narrative and the actors in it, which would not be the case in a still photograph" (Salt, 1992, 85). I will revisit Salt's comparison with still photography later as it becomes relevant to the incidence and function of aerial views in visual media. For now, a few illustrations of high- and low-angle shots in early film reveal a distinction that is highly relevant to the main argument that this essay proposes: Point of View shots with the camera assuming a character's gaze seem to be considerably more frequent than Objective shots that only the cameraman can legitimately appropriate. The Point of View shot obviously designates a diegetic feature of the visual narrative whereas the Objective shot maintains its meaning independent of the diegesis and is may be labeled as a non-diegetic construct.

In analyzing the continuity of action across cuts, Barry Salt (2009, p. 58) remarks: "The consolidation of Williamson's methods of film construction was carried out by other

British film-makers in 1903. The first of these was *Daring Daylight Burglary*, made by the Mottershaws at the Sheffield Photographic Company at the beginning of the year. The film starts with an onlooker leaving the high-angle first shot of a burglar breaking into the back of a house and running off into the next shot of a street elsewhere in which he alerts the police”.¹⁰ The shot (illustrated in Figure 24 below) shows only a slight elevation of the camera from the customary horizontal eye-level, but it may be the first incidence of parting with the horizontal axis of filming especially because it is an instance of the rarer Objective shot independent of any character’s Point of View.



Figure 24. *A Daring Daylight Burglary* (1903) – Two frames showing High Angle Camera Shots.

The unusual low- and high-angle shots in early film were used almost exclusively to follow the Point of View of a character. The Objective high- or low-angle shots exist, but are indeed a rarity. The 1910 Vitagraph film *Back to Nature; or The Best man Wins* uses low- and high-angle shots to show people in a lifeboat looking up at people in an ocean

¹⁰As Salt argues, the eye-level shot was the default for early film: “In the previous period [before 1913] we find the occasional rare use of a just slightly depressed or elevated lens angle, but these were always in shots taken at a considerable distance from the actors, and arose out of the nature of the specific location that was being used. [...] The opposite kind of high angle shot made more or less necessary by the surroundings of the scene, or alternatively done to show clearly what is going on, is also used on rare occasions as in a high angle shot of horses in the 1907 Pathe *Voyous de l’ouest* and the shot of Brutus’ funeral pyre in the Vitagraph *Julius Caesar* (1908). [...] Such shots fall under the concept of ‘cinematographic angle’, which I owe to Jean Mitry. This denotes those types of compositions and framings which did not and could not occur in the still photography of this period and earlier.”

liner and vice versa. A different high angle shot in the same film, however, shows an objective view of the ocean liner's deck, a view that is not preceded by any indication of someone looking down at the deck. Rare as they were, the high- and low-angle shots were too costly to not have a function. But for the early times of film, the function of these shots seems to be peculiar to each instance without insinuating a generic across-the-board meaning. Some of the high-angle shots are perspectival and thus correlated with large shot scales to show panoramas, literally putting things in perspective. Others show groups of people in an effort of increasing the amount of information provided in one single shot.

It is worth noting that the high- and low-angle shots in the early film cover a rather reduced range of the spectrum of possible camera angles: the true bird's eye and worm's eye views are still out of reach. Almost literally out of reach, if we think of the photographic film that must be reached by a certain intensity of light for it to record an impression. Around 1917, Kodak actually produced this potent film. With Kodak getting involved in the defense industry through aviation photography around 1917, military endeavors and aerial shots of landscapes developed a strong bond. It is worth noting that the same contribution of Kodak technology benefited more domestic domains as well, such as urban planning, but somehow the military surveillance aerial shots seem to be a more accessible imagery than strategizing urban highways.

Aerial shots in feature film deserve a closer look not only with respect to filmmaking techniques, but more so with respect to the viewer. How does the viewer process the

vast, rather nondescript space provided by aerial images? What is the rhetoric of aerial landscape? Can we develop a hermeneutics of aerial landscape in cinematic discourse?

2.3. Real Lands, Film Lands, Flatlands: Filming the Middle East; a case study

The world is not unfolding in front of our steady eyes and heads like a movie on a screen. As Gibson (1979) insisted, we, as information seeking entities, encode our environment by moving our eyes and our heads and connecting this visual input with other consistent and contingent sensorial and environmental inputs. Defining our position in relation to other objects and ‘mapping’ the space we are in is thus made possible by our ability to move and record snippets of our surroundings. “Could it be otherwise? After all, we did not develop as god-like observers of the world, but as participants in its life-and-death struggles”(Anderson, 1998, p. 104).

Are then ‘god-like’ views of filmic space perceived as unnatural when depicted on the cinema screen? A more relevant question would address the viewers’ inferential and interpretational process when presented with god-like aerial view. Although Burch (1973) claimed that the off-screen space is more complex due to its varied nature, the on-screen space viewed from above may challenge the simplicity attributed to its putative ‘what’s on screen’ direct accessibility. What does the viewer make of an aerial shots anyway? Are these actually processed as spaces or do they jump into a higher level of signification by alluding to possible inter-cultural hierarchies? It all depends on

how they are employed in the filmmaking endeavors. And this nature of extreme views from above and their dialectics are exactly the issue addressed in the discussion below. Ian Watt (1960) contends: "I take it that whereas explanation, from *explanare*, suggests a mere making plain by spreading out, explication, from *explicare*, implies a progressive unfolding of a series of literary implications [...]." The argument may be the Western filming camera that looks at the Middle East simply explains; and indeed it does: it lays the land flat by spreading it out and smoothing it a bit, too.

2.2.1. The Flatlands down below: Dialectics of camera angle in films of the Middle East; An *explanatory* case study

Both increasingly present in the film industry and undergoing continuous debates over space, the Middle East provides an interesting case study for space and spatial perspectives in film. The distinction between spatial layouts in film of Middle Eastern/Arab production and film *about* Middle East of non-Middle Eastern/Arab production may outline stereotypes regarding the East-West cultural dynamic.

The proposed analysis outlines two apparent trends of showing Middle Eastern landscape and discusses the possible interpretations each trend affords. The point of view assumed by the filmmaker when looking at the Middle East through the filming camera seems to reveal rather reliably a bi-modal distribution of East/West views on the Middle East. First, I will elaborate on the ways in which films of Middle Eastern/Arab production and films of Western/US production represent space. The cognitive aspects underlying the viewers' interpretation of point of view and spatial layouts in these two film traditions are then discussed.

Feature films that deal with the Middle East, whether produced by Middle Eastern filmmakers or Western filmmakers, will inevitably contain landscapes stereotypically associated with the region: desert images, crowded streets, crowded markets, chaotic urban settings. Lina Khatib (2006, 2012) contends that depictions of space in film relating the conflict in the Middle East and the Arab world are reliable tropes in a cinematic rhetoric that delineates two culturally separate trends of contemporary filmmaking: When related to the pervasive conflict almost stereotypically defining the region nowadays, Western producers (especially American/Hollywood) treat the film viewer to vast aerial landscapes, long shots taken from above the horizontal eye-level axis. By contrast, films of Middle Eastern production approach visualizations of space from the ground level. As highlighted in one of the sections above, high angle shots and the film discourse have a rather long history. What then makes the elevated camera shots in film on the Middle East worth our attention? Are they peculiar in any informative way? Khatib contends that the view from “above” intimates superiority and objectifies a subdued (or subdue-able) space - together with its human inhabitants - whereas the view from “below” minimizes the psychological distance between the viewer and the film narrative and ‘levels’ the dialog between them. The former gives us a space that is to be accessed only through the narrow but far-reaching & high-resolution scrutiny of the binoculars. The latter makes the space accessible, even suggesting the possibility of on-foot field trips. Space is inherently cultural, it hosts cultures, it physically marks cultural differences. And as such, any given culture tends to be defined in spatial terms, associated with a bounded piece of land. By analyzing the cinematic techniques of

showing space in multi-cultural film productions, Khatib reveals the dynamics between cultures at conflict.

A body of six US-produced films representing the conflict in the Middle East is compared to a body of six Middle East-produced films with the same thematic interest. Depictions of space in the two collections of films make the comparison term. Systematic differences in representing the Middle Eastern landscape group the two film collections not only along the production lines, but also along the film technique lines. The US produced films tend to look at the Middle East from above through aerial views, satellite views, or simply high camera angles. With a significantly lower incidence of high camera angle shots, the films produced in the Middle East confer a more accessible space, a space in which the viewer can place herself. The distant aerial view that, by its nature, removes the character from the narrative space suggests a position of superiority. The 'on-the-ground' take (or view 'from below', Khatib, 2006) brings the conflict to a recognizable, familiar place.

Zero Dark Thirty (2012)

Intertitles separated by fairly long black screens introduce us to the chronicle of the efforts to capture Osama bin Laden. The very first intertitle promises a 'based-on-true-facts' account and thus sets a documentary tone and imposes an objective approach. The viewer is then invited to a series of torture and interrogation scenes that culminate with the accomplished mission: the capture of bin Laden. Although taking place mostly

in interrogation rooms and in offices, the film displays an impressive array of aerial shots, satellite views, helicopter shots, and other elevated camera takes.

The graph below shows the incidence of eye level, high angle, and low angle shots in Bigelow's (2012) *Zero Dark Thirty*. With only 5.5% high angle shots (of which only 1.9% are actual aerial and satellite views), the argument that such elevated camera takes affect the viewer in any way seems farfetched.

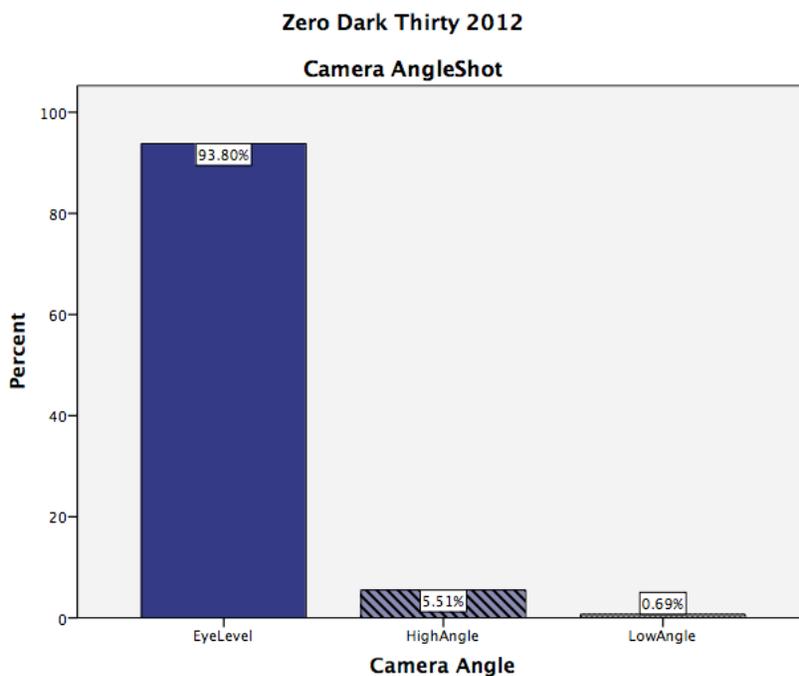


Figure 25. *Zero Dark Thirty* (2010): Percentage of occurrence of high angle, low angle, and eye level film shots.

Duration-wise, the high angle shots take more screen time than the average shot duration of the film. With a 3.5 seconds average shot duration, *Zero Dark Thirty* displays aerial views, satellite maps, and helicopter shots for an average shot length of slightly above 4.5 seconds (see Figure 26 below).

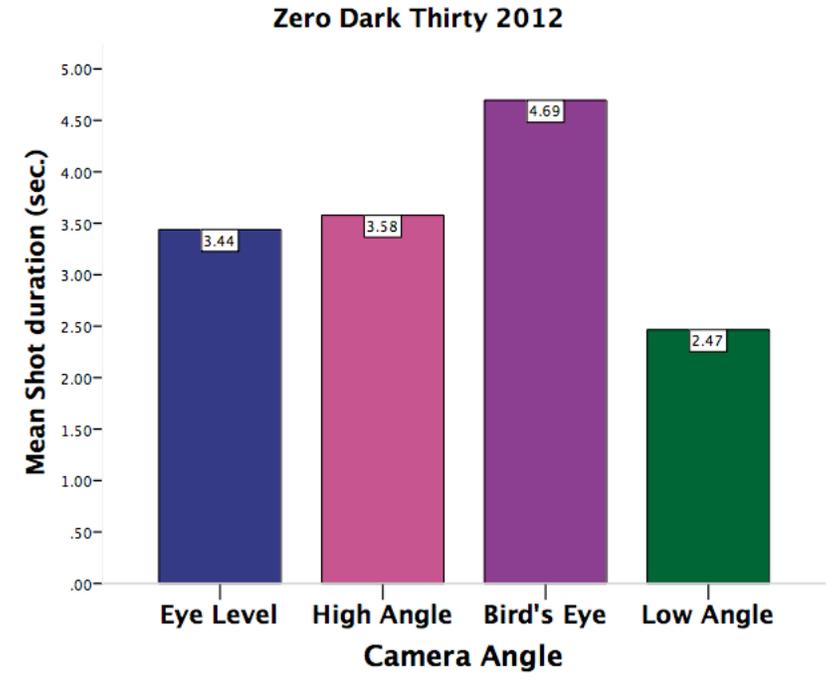


Figure26. *Zero Dark Thirty* (2010): Mean shot duration in three levels of camera angle: Eye level, High Angle, Bird's Eye, and Low Angle.

Research in eye tracking during film viewing indicates that the most 'attractive' screen elements for the viewers' eyes are human faces and high contrast areas. The analyses

on US-produced (or Western-produced) film about the Middle East show rather compellingly that the focal points that generally attract the viewers' eyes are absent in the high angle camera shots. So then, why are these shots used anyway? They are more costly in their technique than the eye-level/ground-level takes and they undermine the very essence of cinematic narrative: inducing a 'being there' sensation in the viewer. Why watch a movie that seems to want to keep the viewer out of the narrative landscape? Why watch a movie that makes the screen visible, a most conspicuous boundary between filmic space and viewers' allocentric space. Good films melt the screen. And with that screen melt-down, the story fills the space and engages the viewer in the 'being there' dynamic. And that is precisely what the movie-goer pays for: The faster/slicker/more inconspicuous the screen melt-down, the better the movie. The high angle shots with their extreme version -- bird's eye view -- impose distance, disengage, ban familiarity and singularity, and trigger objectivity. Bird's eye views and aerial shots of vast landscapes induce the same degree of objectivity that is applied when viewing documentaries. The narrative landscape thus becomes a call for objective, distant interpretation. Bordwell (2008) contends that "Most books introducing narratology start with discussions of the fabula, that spatio-temporal realm in which the action unfolds in chronological order. Then the author goes on to discuss how the world is rendered through patterns of narration – restricted point of view, flashback constructions, and the like. [...] we have access to the fabula only by means of narration. *Narration isn't simply a window through which we watch a preexisting story that we might see from elsewhere*". The *weak filmic space* (mentioned above, see

footnote 3) can be defined as the on-screen space that does not afford narrative information. There is no relationship between the weak filmic space and the narrative space; vast deserts, crowded (as opposed to busy) markets, and traffic-jammed chaotic streets are autonomous entities in the architecture of the film for they can be placed anywhere in the narrative without creating noticeable alterations. They are void of narrative content. And in this case, narration *is* simply a window through which we view a fixed, immutable, distant space.

Concluding remarks

The ideal goal of the research endeavors outlined above was to capture the dynamics of the interaction between the continuity of perception and the discrete nature of event segmentation. The continuity-discreteness dynamic create a loop of, first, a continuous input, secondly, the chunking in encodable units, and finally the meshing the units in a qualitatively continuous account, a re-unified story of a sequence of separate events. Narrative, man-crafted worlds, afford various degrees and types of continuity and discontinuity; they are multidimensional and engage all our senses; they can be disjointed on any of their multiple dimensions; and we all are enthralled by stories. For all these reasons, the narrative became the best candidate for the experimental stimulus.

Events have been defined as spatially- and temporally-contained meaningful units, units that have identifiable, fairly unambiguous boundaries. And as such, they have been

studied as analogs of objects. If one attempts to describe the surrounding environment, one will most likely list the objects or physical units that behave as a coherent composite when displaced. Events should behave the same: they should maintain their coherence if displaced. Now imagine a virtual reality in which parts of individual objects are not adjacent, separated by spatial (or other) gaps. If in that reality one sees disparate tree branches, would one still report seeing a tree? Would that still be a single object to obey the same physical laws an object normally does?

With these questions in mind, I addressed three issues:

1. Narrative in different media: the nature of the event segmentation in literary and cinematic narrative
2. Complex plot constructions: the nature of event segmentation in narratives that break the space/time continuity.
3. Space and the culture boundary: the nature of event segmentation of different spatial depictions in multi-cultural film.

1. Whether written on the page or projected onto the screen, narratives engage the eye. The literary narrative is implicitly visual for those readers who engage in imagery during the process of reading (most of us, but there are exceptions); and it is explicitly visual for the film viewers. Time, abstract and elusive, is not accessible by the senses in an unmediated way. Space is directly perceived. Given the arbitrariness of language, the

text may encode time better than the film; given its visual nature, the film may accommodate space better than text (Chatman, 1980; Magliano et al., 2001a). Using a situation-change judgment paradigm (Magliano et al., 2001a), participants were asked to segment content-equivalent text and film excerpt pairs. Consistent with the affordance of the narrative medium, participants identified more temporal event boundaries in text and more spatial boundaries in film. The inference to be made here is that participants segment along the most salient dimension that the stimulus affords.

2. Film trailers and two structurally complex films (*Memento*, 2002; *Inception*, 2010) were used as stimuli in these studies. Again, events are spatio-temporally bounded units. This definition implies that space and time are the main dimensions around which an event is organized. “Any editing transition is a potential point of discontinuity in the events represented by a movie. Discontinuity can occur in the location of these events; in the flow of time; and in these events relationship to reality [...] (Messaris, 1994). Discontinuities are, obviously, a question of order. “All measurement [...] is based upon order. Indeed, only through suitable arrangements and groupings can we handle vast quantities of material perceptually or cognitively” (Goodman, 1978, p. 13). How do we segment streams of information that are not well arranged and grouped?

Film trailers as well as the two complex films are complex particularly because they break down spatial contingencies and temporal orders. Parsers are then left with the more superficial (or, better said, more accessible) visual features of the stimuli. In the case of film trailers, parsers segment consistent with the promotional nature of the trailer that holds back the ending of the film to create a curiosity-inducing cliffhanger effect.

Therefore, the resolution, or the final act of the film, is almost absent in parsers' segmentation. As in the case above, parsers base their segmentation on features of the trailer (e.g., black screens as conspicuous separations, edits that violate continuity editing). In parsing *Memento* (2002), the transitions between color segments of film and black & white segments of film account for the highest incidence of identified segments. With no apparent pattern, the parsing of *Inception* (2010) is difficult to interpret. Given the incoherent descriptions of events that a few parsers provided, it may be the case that parsers place situation change boundaries at dream level shifts based on the color encoding the dream level and not on coherent online re-construction of the complex narrative.

3. Film is a cultural artifact and as such it engages our culturally-contingent experience. Goodman, in reference Gombrich's research, claims that "the way we see and depict depends upon and varies with experience, practice interests, and attitudes" (1968, p. 10). To illustrate his statement, Goodman provides an example. "When the first fine Japanese films reached us, Western audiences had some difficulty in determining what emotions the actors were expressing" (1968, p. 48). Although not as dramatic as facial expressions, depictions of space are definitely subject to cultural biases and interpretations. The analysis of a sample of US- and Middle East-produced film revealed a pattern in displaying the space of the other culture: the US-produced film looks at the vastness of the Middle East from above. Data collection with participants from the relevant cultures will provide information regarding segmentation along the spatial dimension: is the view from above more or less salient to the Westerner viewer?

So, what is the landscape of stories? What does it look like? If the structure of the narrative affords strong causal connections between adjacent events, the segmentation of the narrative is more linear and probably faster (although I am not aware of reaction time measurement in the domain of event segmentation). I am consistently and the well-trimmed patches on the ground of Germany when I look down right before landing in Munich. The image serves for approximating a linearly-structured narrative with smooth logical transitions from one event to the next. The image is very different when I observe the landscape of the Middle East or Eastern Europe: Either vast empty deserts ready to be farmed out with any degree of randomness, or irregular patches of a world still in construction are the images that approximate our segmentation of complex films, films not strictly obeying the rules of continuity editing.

BIBLIOGRAPHY

- Anderson, J. D. (1998). *The Reality of Illusion: An Ecological Approach to Cognitive Film Theory*. SIU Press.
- Aristotle. (c1982.). *Aristotle's Poetics* / (1st ed.). New York : Norton,.
- Baggett, P. (1979). Structurally equivalent stories in movie and text and the effect of the medium on recall. *Journal of Verbal Learning and Verbal Behavior*, 18(3), 333–356. [http://doi.org/10.1016/S0022-5371\(79\)90191-9](http://doi.org/10.1016/S0022-5371(79)90191-9)
- Bassili, J. N. (1976). Temporal and spatial contingencies in the perception of social events. *Journal of Personality and Social Psychology*, 33(6), 680–685. <http://doi.org/10.1037/0022-3514.33.6.680>
- Bordwell, D. (1985). *Narration in the fiction film* /. Madison, Wis. : University of Wisconsin Press,.
- Bordwell, D. (2012). *Poetics of Cinema*. Routledge.
- Bordwell, D. (2013). *Narration in the Fiction Film*. Routledge.
- Botz-Bornstein, T. (2011). *Inception and Philosophy: Ideas to Die for*. Open Court Publishing.
- Bower, G. H., & Rinck, M. (2001). Selecting one among many referents in spatial situation models. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27(1), 81.
- Branigan, E. (2013). *Narrative Comprehension and Film*. Routledge.
- Buckland, W. (2009). *Puzzle Films: Complex Storytelling in Contemporary Cinema*. John Wiley & Sons.

- Chatman, S. B. (1980). *Story and discourse: Narrative structure in fiction and film*. Cornell University Press. Retrieved from https://books.google.com/books?hl=en&lr=&id=ewrOp9uPjYUC&oi=fnd&pg=PA9&dq=Chatman,+1980&ots=p4glq0_RjN&sig=g2JQhxPngGz9vDMI6gcqKpYkLF4
- Christiansen, M. H., & MacDonald, M. C. (2009). A Usage-Based Approach to Recursion in Sentence Processing. *Language Learning*, 59(s1), 126–161.
- Cutting, J. E., Brunick, K. L., & Candan, A. (2012). Perceiving event dynamics and parsing Hollywood films. *Journal of Experimental Psychology: Human Perception and Performance*, 38(6), 1476–1490. <http://doi.org/10.1037/a0027737>
- Cutting, J. E., Brunick, K. L., & DeLong, J. E. (2011). How Act Structure Sculpt Shot Lengths and Shot Transitions in Hollywood Film. *Projections*, 5(1), 1–16. <http://doi.org/10.3167/proj.2011.050102>
- Cutting, J., & Iricinschi, C. (2015). Re-Presentations of Space in Hollywood Movies: An Event-Indexing Analysis. *Cognitive Science*, 39(2), 434–456. <http://doi.org/10.1111/cogs.12151>
- David, B., & Kristin, T. (2004). *Film art: an introduction*. New York: McGraw-Hill.
- Elsaesser, T. (1990). *Early cinema: Space frame narrative*. BFI Publishing. Retrieved from <http://dare.uva.nl/en/record/127183>
- Erin Brockovich. (c2000.).
- Field, S. (c1994.). *Screenplay: the foundations of screenwriting* / (3rd ed., expanded ed.). New York: Dell Pub. Co.,.

- Fleiss, J. L. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin*, 76(5), 378–382. <http://doi.org/10.1037/h0031619>
- Fleiss, J. L., Levin, B., & Paik, M. C. (2013). *Statistical Methods for Rates and Proportions*. John Wiley & Sons.
- Gibson, J. J. (1979). The Ecological Approach to Visual. *Perception*.
- Glenberg, A. M., Meyer, M., & Lindem, K. (1987). Mental models contribute to foregrounding during text comprehension. *Journal of Memory and Language*, 26(1), 69–83. [http://doi.org/10.1016/0749-596X\(87\)90063-5](http://doi.org/10.1016/0749-596X(87)90063-5)
- Goodman, N. (1968). *Languages of Art: An Approach to a Theory of Symbols*. Hackett Publishing.
- Goodman, N. (1978). *Ways of Worldmaking*. Hackett Publishing.
- Harvey (film). (2015, June 23). In *Wikipedia, the free encyclopedia*. Retrieved from [https://en.wikipedia.org/w/index.php?title=Harvey_\(film\)&oldid=668301404](https://en.wikipedia.org/w/index.php?title=Harvey_(film)&oldid=668301404)
- Heider, F., & Simmel, M. (1944). An Experimental Study of Apparent Behavior. *The American Journal of Psychology*, 57(2), 243–259. <http://doi.org/10.2307/1416950>
- Inception. (c2010.).
- Johansson, G. (1950). Configurations in event perception. *Uppsala, Sweden: Almqvist*, 8.
- Jansson, G., Bergström, S. S., Epstein, W., & Johansson, G. (Eds.). (1994). *Perceiving events and objects*. Hillsdale, N.J: L. Erlbaum Associates.
- Jensen, C. S. (2014). Reduced narration, intensified emotion: the film trailer. *Projections: The Journal for Movies and Mind*, 8(1), 105+.

- Kernan, L. (2009). *Coming Attractions: Reading American Movie Trailers*. University of Texas Press.
- Khatib, L. (2006). *Filming the Modern Middle East: Politics in the Cinemas of Hollywood and the Arab World*. I.B.Tauris.
- Khatib, L. (2012). *Image Politics in the Middle East: The Role of the Visual in Political Struggle*. I.B.Tauris.
- Kim, J. (1969). Events and Their Descriptions: Some Considerations. In N. Rescher (Ed.), *Essays in Honor of Carl G. Hempel* (pp. 198–215). Springer Netherlands. Retrieved from http://link.springer.com/chapter/10.1007/978-94-017-1466-2_10
- Koffka, K. (1935). *Principles of gestalt psychology* / (First edition.). New York : Harcourt, Brace.
- Köhler, W. (1929). *Gestalt psychology* (Vol. xi). Oxford, England: Liveright.
- Kurby, C. A., & Zacks, J. M. (2008). Segmentation in the perception and memory of events. *Trends in Cognitive Sciences*, 12(2), 72–79. <http://doi.org/10.1016/j.tics.2007.11.004>
- Lamorisse, A. (1957). *The Red Balloon*. Family, Fantasy, Short.
- Magliano, J. P., Miller, J., & Zwaan, R. A. (2001a). Indexing space and time in film understanding. *Applied Cognitive Psychology*, 15(5), 533–545. <http://doi.org/10.1002/acp.724>
- Magliano, J. P., Miller, J., & Zwaan, R. A. (2001b). Indexing space and time in film understanding. *Applied Cognitive Psychology*, 15(5), 533–545. <http://doi.org/10.1002/acp.724>

- Magliano, J. P., & Zacks, J. M. (2011). The Impact of Continuity Editing in Narrative Film on Event Segmentation. *Cognitive Science*, 35(8), 1489–1517.
<http://doi.org/10.1111/j.1551-6709.2011.01202.x>
- Maier, C. D. (2009). Visual evaluation in film trailers. *Visual Communication*, 8(2), 159–180. <http://doi.org/10.1177/1470357209102112>
- Maier, C. D. (2011). Structure and Function in the Generic Staging of Film Trailers: A Multimodal Analysis. *Pragmatics and Beyond New Series*, 211. Retrieved from <http://forskningsbasen.deff.dk/Share.external?sp=S929bdad0-06c0-11e0-a92b-000ea68e967b&sp=Sau>
- Memento. (2002).
- Messaris, P. (1994). *Visual "literacy": Image, mind, and reality* (Vol. xii). Boulder, CO, US: Westview Press.
- Michotte, A. (1963). *The perception of causality* (Vol. xxii). Oxford, England: Basic Books.
- Morrow, D. G., Bower, G. H., & Greenspan, S. L. (1989). Updating situation models during narrative comprehension. *Journal of Memory and Language*, 28(3), 292–312. [http://doi.org/10.1016/0749-596X\(89\)90035-1](http://doi.org/10.1016/0749-596X(89)90035-1)
- Newtonson, D. (1976). Foundations of attribution: The perception of ongoing behavior. *New Directions in Attribution Research*, 1, 223–247.
- Pronin, E., Jacobs, E., & Wegner, D. M. (2008). Psychological effects of thought acceleration. *Emotion*, 8(5), 597–612. <http://doi.org/10.1037/a0013268>

- Radvansky, G. A., & Copeland, D. E. (2006). Walking through doorways causes forgetting: Situation models and experienced space. *Memory & Cognition*, *34*(5), 1150–1156.
- Radvansky, G. A., & Zacks, J. M. (2014). *Event Cognition*. Oxford University Press.
- Rinck, M., & Bower, G. H. (2000). Temporal and spatial distance in situation models. *Memory & Cognition*, *28*(8), 1310–1320.
- Ryan, M.-L. (2004). *Narrative Across Media: The Languages of Storytelling*. U of Nebraska Press.
- Salt, B. (2009). *Film Style and Technology: History and Analysis*. Starword.
- Smith, T. J. (2012). The Attentional Theory of Cinematic Continuity. *Projections*, *6*(1), 1–27. <http://doi.org/10.3167/proj.2012.060102>
- Suckfüll, M., & Moellering, K. (2015). The differential success of movie trailers. *Journal of Retailing and Consumer Services*, *22*, 138–144.
- Swallow, K. M., Barch, D. M., Head, D., Maley, C. J., Holder, D., & Zacks, J. M. (2010a). Changes in Events Alter How People Remember Recent Information. *Journal of Cognitive Neuroscience*, *23*(5), 1052–1064. <http://doi.org/10.1162/jocn.2010.21524>
- Swallow, K. M., Barch, D. M., Head, D., Maley, C. J., Holder, D., & Zacks, J. M. (2010b). Changes in Events Alter How People Remember Recent Information. *Journal of Cognitive Neuroscience*, *23*(5), 1052–1064. <http://doi.org/10.1162/jocn.2010.21524>

- Thompson, K. (1999). *Storytelling in the new Hollywood: understanding classical narrative technique* /. Cambridge, Mass. : Harvard University Press,.
- Tversky, B., Zacks, J. M., & Hard, B. M. (2008). The structure of experience. *Understanding Events: How Humans See, Represent, and Act on Events*, 436–464.
- Tversky, J. B., Zacks, J. B. M. M., & Hard, B. M. (2010). Talking About Events. In J. Bohnemeyer & E. Pederson (Eds.), *Event Representation in Language and Cognition*. Cambridge University Press.
- Wittgenstein, L. (c1997.). *Philosophical investigations* / (2nd ed., reissued German-English ed., reissued 2nd ed.). Oxford, UK ; Blackwell,.
- Zacks, J. M., Speer, N. K., & Reynolds, J. R. (2009). Segmentation in reading and film comprehension. *Journal of Experimental Psychology: General*, 138(2), 307–327. <http://doi.org/10.1037/a0015305>
- Zacks, J. M., Speer, N. K., Swallow, K. M., Braver, T. S., & Reynolds, J. R. (2007). Event perception: A mind-brain perspective. *Psychological Bulletin*, 133(2), 273–293. <http://doi.org/10.1037/0033-2909.133.2.273>
- Zacks, J. M., & Swallow, K. M. (2007). Event Segmentation. *Current Directions in Psychological Science*, 16(2), 80–84.
- Zacks, J. M., & Tversky, B. (2001). Event structure in perception and conception. *Psychological Bulletin*, 127(1), 3–21. <http://doi.org/10.1037/0033-2909.127.1.3>

Zwaan, R. A., Langston, M. C., & Graesser, A. C. (1995). The Construction of Situation Models in Narrative Comprehension: An Event-Indexing Model. *Psychological Science*, 6(5), 292–297.

Zwaan, R. A., Magliano, J. P., & Graesser, A. C. (1995). Dimensions of situation model construction in narrative comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21(2), 386–397. <http://doi.org/10.1037/0278-7393.21.2.386>

Zwaan, R. A., Radvansky, G. A., Hilliard, A. E., & Curiel, J. M. (1998). Constructing Multidimensional Situation Models During Reading. *Scientific Studies of Reading*, 2(3), 199–220.

FILMOGRAPHY

Beck, J. (Producer) & Koster, H. (Director). (1950). *Harvey*. United States: Universal Studios Home Video. *Brunetti, D., & Chafin, C. (Producers) & Fincher, D. (2010). *The Social Network*. United States: Columbia TriStar Home Entertainment.

Cruise, T., & Wagner, P. (Producers), & Woo, J. (Director). (2000). *Mission: Impossible II*. United States/ Germany: Paramount Home Entertainment.

Gilbert, B.(Producer), & Higgins, C. (1980). *Nine to Five*. United States: Twentieth Century Fox Home Entertainment.

Gordon, C., & Gordon, L. (Producers), & Harlin, R. (Director). (1990). *Die Hard 2*. United States: Twentieth Century Fox Home Entertainment.

Hecht, H., & Ross, F. (Producers), & Tourneur, J. (Director). (1950). *The Flame and the Arrow*. United States: Warner Home Video. *Hughes, J. (Producer), & Columbus, C. (Director). (1990). *Home Alone*. United States: Twentieth Century Fox Home Entertainment.

Karz, M., Rice, W. A., & Rosen, J. (Producers), & Marshall, G. (Director). (2010). *Valentine's Day*. United States: Warner Home Video.

Kramer, S. (Producer & Director). (1960). *Inherit the Wind*. United States: MGM Home Entertainment.

Lamorisse, A. (Director). (1956). *The red balloon*. France: Films Montsouris

Lucas, G. (Producer), & Kershner, I. (Director). (1980). *Star Wars: Episode V - The Empire Strikes Back*. United States: Twentieth Century Fox Home Entertainment.

Mankiewicz, J. L. (Producer) & Cukor, G. (Director) (1940). *The Philadelphia Story*. United States: Warner Home Video.

Meyers, N. (Producer & Director). (2000). *What Women Want*. United States: Paramount Home Video.

Milestone, L. (Producer & Director). (1960). *Ocean's Eleven*. United States: Warner Home Video.

Nolan, C. (Producer & Director). (2010). *Inception*. United States/United Kingdom: Warner Home Video.

Nolan, C. (2000). Memento [film]. *Newmarket Capital Group, Team Todd, I Remember Productions, Summit Entertainment, USA*.

Parker, A., Jones, L., Rudin, S., Brown, D., Watson, E., Carlyle, R., Williams, J., & McCourt, F. (2000). *Angela's ashes*. Hollywood, Calif.: Paramount.

Rafelson, B. (Producer & Director). (1970). *Five Easy Pieces*. United States: Columbia TriStar Home Entertainment.

Schwary, R. L. (Producer), & Redford, R. (Director). (1980). *Ordinary People*. United States: Paramount Home Video.

Winkler, I. (Producer), & Scorsese, M. (Director). (1990). *Goodfellas*. United States: Warner Home Video.

Zanuck, D. F. (Producer) & Ford, J. (Director). (1940). *The Grapes of Wrath*. United States: Twentieth Century Fox Home Entertainment.