Event segmentation and seven types of narrative discontinuity in popular movies

James E. Cutting *
Cornell University, USA

A R T I C L E   I N F O

Article history:
Received 23 September 2013
Received in revised form 10 March 2014
Accepted 15 March 2014
Available online xxxx

PsycINFO classification:
2323, visual perception
2340, cognitive processes
2610, literature and the fine arts

Keywords:
Events
Film style
Movies
Narrative
Scenes
Segmentation

A B S T R A C T

Using a sample of 24 movies I investigate narrative shifts in location, characters, and time frame that do and do not align with viewer segmentations of events (scenes and subscenes) in popular movies. Taken independently these dimensions create eight categories, seven of change and one of nonchange. Data show that the more dimensions that are changed the more viewers agree on their segmentations, although the nonadditive variations across the seven change types are large and systematic. Dissolves aid segmentation but over the last 70 years they have been used less and less by filmmakers, except for two infrequent shift types. Locations and characters are strongly yoked, jointly accounting for most narrative shifts. There are also interactions of shift types over the 70-year span and across genres, as well as differences that affect the scale of the establishing shot in a new scene. In addition, several aspects of the narratives of individual movies affect the distributions of shift types. Together these results suggest that there are at least four different signatures of narrative shifts to be found in popular movies—general patterns across time, patterns of historical change, genre-specific patterns, and film-specific patterns.

© 2014 The Author. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/3.0/).

1. Introduction: popular movies, continuity, and discontinuity

In drama, scene refers to a division within an act of a play, indicated by a change of locale, abrupt shift in time, or the entrance or exit of a major character. (Polking, 1990, p. 405)

Works of theater, literature, opera, and film typically present stories that are broken into chunks. A scene is a medium-size chunk found in all the narrative arts and often synonymous with the concept of an event. It is the job of the film director and editor not only to present these as separate events, but also to elide them so that a semblance of continuity is maintained. For example, in movies scene boundaries are often covered by nondiegetic (background) music or occasionally by a voiceover to maintain the semblance of seamlessness. Continuity is particularly critical to popular movies because they are broken further into shots—runs of successive frames divided most often by cuts and abrupt changes in camera position. In this article I focus on discontinuities across the shot boundaries of the visual narrative that do and do not signal a new event.

Continuity, long an idea central to filmmaking, is a hybrid concept. It is partly psychological; continuity is what drives a movie smoothly to a climax, keeping the viewer engaged and anticipating possible outcomes while ingesting new information from the plot (Smith, 2012). But continuity is also a physical concept and can be deliberately broken along three basic and typically salient parameters of a scene—location, character, and time. As implied by Polking in the epigram, scenes canonically take place in a single location, with a single set of characters, during a single time frame. However, partly because the job of filmmakers is to mask the sutures across changes in these parameters, the definition of a scene, like that of continuity, is not only simply structural but also psychological. It is as much defined by film viewers as by filmmakers. My tasks in this article are to look at the concordance of viewer segmentations across the various types of narrative shifts in a representative sample of popular movies; to catalog the relative frequencies of the shifts in location, characters, and time; and to suggest psychological causes for the differences that affect segmentation; and to reveal under-appreciated aspects of Hollywood style.
2. Location, character, and time shifts across scenes

Messaris (1994) surveyed the nature of shifts across edited scenes. From nine fictional television programs and a handful of movies, he cataloged some possibilities. He counted shifts in location and shifts in time (flashbacks and fast-forward), but he also considered shifts in “reality” — moves from diegetic (narrative) time to dreams, memories, or fantasies, and back again. I will denote such changes as additional shifts in time (see also Cutting & Iricinschi, in press). The reason is that visualized dreams occur during sleep (so the character is projected to a different narrative time), visualized memories are of previous events (so reflect back in time), and visualized fantasies occur in time frames that have never happened (and are therefore removed from the diegetic time stream of events). Moreover, such “reality” shifts are relatively rare in movies. Indeed, they occur in only three of 24 movies in the sample I will discuss, Nine to Five (1980), Ordinary People (1980), and Inception (2010), and all of these movies use special stylistic tools to denote such shifts.

In this article, I follow Messaris’ lead in cataloging shifts of location and time, but I also include character shifts as well. I consider a character shift to occur either by addition, subtraction, or complete change of major characters across shots. Cutting and Iricinschi (in press) found that most movies have ten or fewer major characters and that each typically appears in at least 8% of all scenes. Thus, the character ensemble is not usually difficult to track. Individual characters are also typically invariant; that is, they usually do not change much physically, perhaps beyond aging, across the length of a movie.

Time shifts, on the other hand, are often more difficult to detect. To be sure, many are obvious — day turns into night across a cut; blistering heat in one shot is replaced by a downpour in the next; a person runs through an airport and then a plane takes off; or two shots are separated by a dissolve or a fade, the traditional methods of denoting the passage of time (Cutting, Brunick, & Delong, 2011). Unless marked in one of these visual or conceptual ways, however, it is often impossible to tell if time has shifted. Without evidence to the contrary, therefore, my students and I have coded scene changes as not having time shifts.

Moreover, many temporal ellipses seem insufficiently important or are simply too small to create a narrative shift. Consider a three-shot sequence about 20 min into Erin Brockovich (2000). In the first Erin (Julia Roberts) sits at her desk as a new employee in a law office trying to understand connections among folders for a case. A dissolve then occurs into the second shot and Erin is still at her desk with the same materials but seen from a different angle. A change of lighting suggests that it is later in the day. She then starts to get up and a cut occurs. The third shot reveals her just arriving at the desk of a co-worker and asking a question. The diegetic (narrative) time gap between the first and second shots is likely to be several hours and yet no important action has happened. The dissolve and the juxtaposition of the two shots simply denote that she has been working throughout. The time between the end of the second and the beginning of third shot is probably 2 or 3 s, too trivial to count as a time shift. I call this latter time deletion a jump shift. In playing this sequence to students I have found that viewers agree that the considerably different-sized temporal lacunas between the first and second and between second and third shots do not signal a change of scene.

Similarly, small or large time shifts can occur within what Metz (1974) called a syntagma — a scene consisting of short subunits, often single shots, that have no direct relation with one another but that build to make a point. Consider two examples. First, 11 min into Erin Brockovich there is a parallel syntagma — a series of shots that alternate versions of a repeated theme to make a point. Erin collects newspapers (two shots), circles job ads (four shots), and serially calls potential employers (19 shots). The whole scene takes about 1 min of screen time. Traditional continuity is purposefully violated. For the phone calls, she is alternately dressed in three different blouses, she has three different demeanors for job inquiry, and she calls from both inside her house and from a pay phone, sometimes with and without her baby. Clearly such action takes place in multiple locations and at multiple times perhaps over weeks, but it coheres as a single, distributed event about a character’s difficulties finding a job. Moreover, when I have played this sequence to viewers they judge it to be a single scene, often without noticing anything peculiar about it.

Second, almost 30 min into Ordinary People there is a party sequence consisting of 18 shots that forms a bracket syntagma. Its shots effectively describe the interactions of upper-middle class adults, showing snippets of different conversations on unrelated topics that take place in different rooms of a house among different characters. One can guess that the sequence covers the span of about an hour of diegetic time. Thus, although time cannot be continuous, perhaps not even serially ordered, across the shots they nonetheless cohere as a single scene and viewers judge them to be so.

More than character or time shifts, location shifts can be hard to define. Separate locations are most often delimited by doors (Bordwell & Thompson, 2004; Burch, 1973; see also Radavsky & Copeland, 2006) or other barriers, physical or psychological. Many scenes are actions that take place within their confines, but naturalistic outdoor scenes can be unbounded (Magliano, Miller, & Zwaan, 2001). Moreover, if actors move through a space, even through doors, and the action and content are not broken then the entirety can be considered a single location. Magliano et al. (2001) and Cutting and Iricinschi (in press) have worked through many possibilities and give more examples of what might count as a single location, and what might not. As with jump shifts, which allow short bits of time to be elided out of the scene, small changes in location — from one side of a room to another, from inside a car looking out to outside a car looking in, and even from one end of a telephone call to the other end (which may be a world away) — may not matter to the viewer. Psychologically important cinematic devices like eye-line matches, shot/reverse-shot composition, and point of view editing help knit disparate spaces together.

3. Key terms and concepts

Again, a cut is a sharp transition between shots, one frame to the next. A dissolve is an optical mixture of two shots over a dozen to several dozen frames (assuming 24 frames/s), where the first shot dominates at the beginning of the dissolve, and is gradually replaced until only the second shot remains. A fade is like a dissolve although it goes through black between the two shots and is usually about twice as long or longer. A wipe is a transition in which a second shot replaces a first along

---

1 McCloud (1991) also presented an impressive analysis of the transitions across panels of comics that are similar to those of Messaris for film. See also Cohn (2013).
2 Of course, there are occasional exceptions. In Mission: Impossible II (2000) the artifacts of wearing tight-fitting masks and external throat devices that change voices allow protagonist Ethan Hunt (Tom Cruise) and antagonist Sean Ambrose (Dougray Scott) to impersonate other people. But such examples are rare exceptions.
3 I distinguish a jump shift from a jump cut. In a jump cut the camera occupies the same or nearly the same position in two consecutive shots, but the time frame has jumped ahead from the end of the first shot to the beginning of the second. In a jump shift the camera occupies different positions in the two shots, but the same amount of time may have been deleted. The purpose of this second kind of cut is to move the action forward, omitting needless detail about how a character got from one position to the next. Traditionally, a jump cut was thought to violate continuity; a jump shift does not. In fact it seems to promote an intensified continuity (Bordwell, 2006). However, jump cuts seem to have become more and more acceptable, and are quite common in contemporary television.

---

4 Eye-line matches across shots knit the line of gaze of one individual to another in conversation. Shot/reverse-shot alternation is the most common form of conversational scene construction, and typically one person is displaced slightly to the right and looks left, while the other person is displaced slightly left and looks right. This construction typically creates eye-line matches. Point-of-view (POV) edits, where a character looks off-screen and the next shot reveals what that character is looking at, are sometimes also called eye-line matches. POV edits seem to mimic the psychological principle of joint attention (see for example Carpenter, Nagell, & Tomasetto, 1998), where when the character looks off screen we follow his or her gaze in order to understand what is being looked at.
a boundary that traverses the screen, usually left to right. Collectively, I will call fades, dissolves, and wipes by the term noncuts.

Together, I will call cuts and noncuts shot transitions. These contrast with the change from one potential scene to the next, which I will call narrative shifts. Narrative shifts occur with visually or conceptually discernible changes of location, character, or time. Scene changes are psychologically determined by viewers, generally following the narrative shift information. Notice also that there are typically many shot transitions within a scene. However, when scenes change from one shot to the next there is necessarily also a shot transition between them, and almost always a narrative shift of some kind. The nature of the shot transitions and narrative shifts are used in this article to inform viewer segmentations of scenes (events).

Shot scale is the measure of the size of a character or some similar sized object within the frame of the image. Cutting, Brunick, and Candan (2012) found it to be the most potent predictor of event segmentation in popular movies. By convention (see Bordwell & Thompson, 2004; Salt, 2006) shot scale is divided into seven categories: (1) extreme long shot, (2) long shot, (3) medium long shot, (4) medium shot, (5) medium close up, (6) close up, and (7) extreme close up. Fig. 1 shows four of these in-stills from four successive shots about 12 min into Goodfellas (1990). The upper left shows an extreme long shot (1), where the vertical extent of the characters soon to be seen is considerably smaller than the vertical extent of the frame. The upper right follows in a medium-long shot (3, here a mid-thigh up view) of a trucker about to be involved with Jimmy Conway (Robert De Niro), a member of an organized crime group. The lower left shows an extreme close up (7) of a wallet and fake driver’s license held in Jimmy’s hand, and the lower right a medium close up (5, here mid-chest to shoulders and up) of Jimmy. Other shot scales are in between these, with a long shot (2) barely encompassing the head and feet of a character, a medium shot (4) showing him or her from the waist or stomach up, and a close up (6) showing just the face and the top of the shoulders.

Scenes are typically tied to the physical measures of narrative shifts. But consider once more the syntagmas from Erin Brockovich and Ordinary People. Continuity of location is violated in the former, continuity of characters in the latter, and continuity of time is violated in both. Thus, discontinuities along one or more of Polking’s attributes – location, characters, and time – do not inherently imply that the scene has changed. Such changes are typically necessary but not sufficient.5 The lesson, again, is that scene integrity and scene shifts are psychophysical hybrids; they are characterized as much by their psychological impression as by their physical construction. An idea adapted from Aristotle’s Poetics (Fyfe, 1932) is that scenes, like the larger works they are part of, should have beginnings, middles, and ends. Indeed, Cutting et al. (2012) found that scenes in popular movies have a physical arc. That is, the first shot in a scene, called an establishing shot, is generally longer in duration than those that follow and it is also longer in shot scale. The final shot of a scene is also generally longer in duration than those that precede it (except for the first).

Subscenes are parts of scenes or sequences, and they are a necessary descriptive unit in the segmentation of movies (see Bellour, 1976). A major reason for this unit is the narrative tactic of parallel action, used particularly in action and adventure movies, of cutting back and forth between ongoing actions of a protagonist and an antagonist. The protagonist will do something in one location, the antagonist something in another, and this alternation may continue for some time. Such units are subscenes within the larger unit (typically called a sequence) where the two characters move towards their eventual confrontation.

Although it is important to allow for both scenes and subscenes, this is not the place to try to distinguish them further. Instead, I will combine them in a single analysis. The importance of this combination is that it ties both units to the psychological concept of an event — an action by someone (or a group) in a given place and at a given time. Events have proved important in perception, language understanding, literature, and movies (Cutting, 1981; Cutting et al., 2012; Zacks & Magliano, 2011; Zacks, Speer, & Reynolds, 2009; Zacks, Speer, Swallow, & Maley, 2010; Zwaan, Langston, & Graesser, 1995; Zwaan, Magliano, & Graesser, 1995; Zwaan & Radavsky, 1998). Moreover, events can also be of different sizes and nested within one another (Shaw & Cutting, 1980).

Finally, film style is the particular combination of storytelling techniques used in a given movie. These include variations in shot duration, shot scale, shot transitions, camera positions and movement, staging, lighting, color, and so forth. Hollywood style is the familiar combination of a large collection of such techniques (Bordwell, Staiger, & Thompson, 1985).

4. The movie corpus, previous results, and a taxonomy of narrative shifts

In previous research my students and I selected 160 popular, English language movies, ten each released five years apart from 1935 to 2010, spread across five genres. All were among the highest grossing movies of their release year. In search of long-term trends in film style, we segmented these movies into their individual shots and investigated various shot properties (Cutting, Brunick, DeLong, Iricinski, & Candan, 2011; Cutting, DeLong, & Nothelfer, 2010). From this larger set of movies we selected 24 — one drama, one comedy, and one action movie released ten years apart from 1940 to 2010 (Cutting et al., 2012).6 We then had eight viewers, three per movie, view each movie twice. On a second pass through a given movie, each viewer segmented the shots into events. Five viewers were naive to the project, although their results were not systematically different than those of the other three. Each viewer was assigned their movie as they had the time to view and segment them. We provided no strict definition of an event to these viewers, yet without consultation they agreed in their segmentations 91% of the time (median κ = .56).

Cutting and Iricinski (in press) then analyzed narrative shifts according to whether they presented changes to a new location, character, or time frame, or an old one previously seen in the movie. We found that scale and durations were greater for shots introducing new locations than old ones. In this article I focus on aspects of viewer segmentations not previously analyzed by Cutting et al. (2012) or by Cutting and Iricinski (in press). That is, I outline seven classes of narrative shifts, how each of the viewers of each film used them to segment a movie, and how the shifts follow from the genre and the structure the visual narrative.

The more than 31,000 shot boundaries across the 24 movies were divided into eight categories, which I will designate with a three-digit nomenclature, also given in Table 1. The types are: 111, 110, 101, 100, 011, 010, 001, and 000, where the first digit corresponds to the presence (1) or absence (0) of a location shift, the second to the presence or absence of a character shift, and the third to the presence or absence of a time shift. The last category – 000 – is a null-shift. It is used here only as a basis for signal detection analyses. The remainder of the nearly 3000 scene and subscene changes is distributed across the seven types of narrative shift. In what follows I focus first on viewer segmentations and their correlates in Hollywood style, and second on the structure of Hollywood style in light of the seven types of narrative shifts.
5. Event segmentation and Hollywood style

5.1. Concordance of viewer segmentations

How well do observers agree in their segmentations? For each movie, I first divided them across the seven categories of location, character, or time shift for each movie, and then separated them by whether the shift was preceded by a cut or by a noncut. Sensitivity ($d'$) analyses were performed using the group hit rates for each of the seven classes of narrative shift for both cuts and noncuts within a movie, then averaging across them. Within a movie, the pooled false alarm rates were determined from segmentations at a null shift (000). Fig. 2 shows the results, with dark bars indicating those scenes and subscenes separated by cuts, lighter bars indicating those separated by noncuts.

Notice first that there are large differences across shift types. Consider the combined results involving those preceded by cuts (87% of the sample) and by noncuts. The sensitivity to shifts in which location, characters, and time all change (111) was statistically greater than mean of those in which only two of these dimensions changed (110, 101, and 011; $d' = 3.06$ vs. 2.60, matched-pairs $t(22) = 3.45, p < .003$, Cohen's $d = 1.47$). Moreover, the values for these two-dimension changes were also statistically greater than the mean of those shifts along only one dimension (100, 010, and 001; $d' = 2.60$ vs. 2.22, $t(23) = 4.08, p < .0005, d = 1.70$). This result – that changes over more dimensions make segmentation easier – seems reasonable. However, it does not inherently suggest additivity of scene-change information. As we will see, these dimensions interact in various ways.

5.2. The effects of shot transitions on detecting scene changes

As should also be clear from Fig. 2 there was greater observer efficacy in identifying scene boundaries in the presence of noncuts than cuts (mean $d' = 2.78$ vs. 2.38; matched-pairs $t(15) = 3.77, p < .002$, Cohen's $d = 1.95$), and this was true for two of the four individual comparisons for which there are sufficient data (all from older movies): those shifts with location, characters, and time changing (111; matched-pairs $t(7) = 2.79, p < .03, d = 2.11$), those with location and time changing but not characters (101; $t(8) = 3.55, p < .009, d = 2.68$).

Thus dissolves, fades, and wipes aid segmentation as filmmakers have intended. This result is not surprising, but perhaps more importantly the data also support the idea that contemporary viewers can segment movies quite well from cuts alone. Moreover, the decline in $d'$ values can be interpreted as a success of newer Hollywood style by enhancing the continuity of movies. Given that dissolves, fades, and wipes can take up valuable screen time, are not wholly required, and can slow or stop the intended flow of a movie, contemporary filmmakers can afford not to use them (Cutting, Brunick, & DeLong, 2011). Nonetheless, there are some residual effects of noncuts that I will discuss below.

In contemporary cinema cuts comprise almost 99% of all shot transitions and even 87% of those separating scenes. But this has not always been the case. Film style has changed. At the end of the 19th century, the first movies were single shots of single scenes, soon to be followed by movies with several shots, each one a separate scene and each separated by a dissolve or a fade. Dissolves and fades were used particularly to separate scenes that differed in time. On the other hand, wipes, which developed later and which have never been as common as the other two, were used typically to change locations and characters without changing time. Straight cuts appeared late in the first decade of film, but they were always used within a scene. Cuts did not dominate scene changes until the 1960s (Carey, 1974; Cutting, Brunick, & DeLong, 2011).

In this sample of movies the combined use of noncuts for 111 shifts fell from 90% in the 1940s and 1950s to 10% in most recent decade ($r = −.76, t(22) = 5.53, p < .0001, d = 2.35$), and their use fell almost as strikingly for the other common shifts (110, $r = −.64, p < .0008$; 101, $r = −.56, p < .004$; 100, $r = −.47, p < .02$). However, the marked exceptions are those involving time shifts without location shifts (011 and 001). As noted below, these are the least common shifts, and eight of twelve (67%) of the former and 20 of 28 (71%) of the latter

---

Footnotes:

7. Normally, one does not pool responses across observers for sensitivity ($d'$) analyses since grouped rates tend to inflate $d'$. However, since each observer only contributed one observation per segmentation, it seemed best to consider each narrative shift as the unit of investigation rather than each observer. In addition, the term sensitivity is not quite right in this context, since it emphasizes only the role of the perceiver. A better term might be efficacy, which can stand for both the information about locations, characters, and time in the movie and the viewer’s consideration of them. Finally, false alarms (segmentations to stimuli without location, character, or time shifts) $−.001$ averaged 2.2% across all movies. Values of $d'$ were computed for each movie with sufficient numbers of hits, then averaged across movies.

8. In several cases, one or more films do not have particular transition types and thus are not part of the analysis.
Table 1

<table>
<thead>
<tr>
<th>Type</th>
<th>Description</th>
<th>Percentage of all narrative shifts</th>
<th>Mean scale of the establishing shot</th>
</tr>
</thead>
<tbody>
<tr>
<td>111</td>
<td>Shifts in location, characters, and time</td>
<td>30.3</td>
<td>2.03</td>
</tr>
<tr>
<td>110</td>
<td>Shifts in location and characters, but not time</td>
<td>41.0</td>
<td>2.30</td>
</tr>
<tr>
<td>101</td>
<td>Shifts in location and time, but not characters</td>
<td>6.3</td>
<td>1.07</td>
</tr>
<tr>
<td>100</td>
<td>Shift in location, but not characters or time</td>
<td>7.6</td>
<td>1.94</td>
</tr>
<tr>
<td>011</td>
<td>Shifts in characters and time, but not location</td>
<td>0.4</td>
<td>3.11</td>
</tr>
<tr>
<td>010</td>
<td>Shift in characters, but not location or time</td>
<td>13.5</td>
<td>2.62</td>
</tr>
<tr>
<td>001</td>
<td>Shift in time, but not characters or location</td>
<td>0.9</td>
<td>2.25</td>
</tr>
<tr>
<td>000</td>
<td>No shift in location, characters, or time</td>
<td>–</td>
<td>(3.88)</td>
</tr>
</tbody>
</table>

Codes of Magliano and Zacks (2011).

- a Action discontinuities.
- b Spatial-temporal discontinuities.
- c Continuity edits.

The mean use of noncuts for the other five shifts across release years is 15%, yielding a strong pattern of differences across the seven types ($\chi^2(6) = 2116.6, p < .0001$). Moreover, although there are too few 011 and 001 shifts to calculate statistical trends, there is no apparent decline in the use of dissolves for these over the last 70 years.

Why is the dissolve retained for these two types of shifts? Dmytryk (1984, pp. 83–84) suggested that the purpose of a dissolve is to act like a “time machine” giving full warning to the movie viewer that time has elapsed. I take the high percentages of dissolves for 011 and 001 narrative shifts as evidence that these are otherwise difficult to comprehend.

5.3. How shot scales vary across shift types to aid segmentation

Character shifts are critical to movie narratives, but environments are also important in showing the movements of those characters through the story. If characters are to be recognized they should be relatively large within the frame (implicating shorter shot scales); if locations are to be registered by viewers more of them should be visible (implicating longer shot scales). Movie framing thus creates a conflict. The larger the images of the characters the less of the location can be seen, and the more of the location that can be seen the less distinct are the characters. Thus, narrative shifts involving characters and locations, together and separately, ought to have distinctive shot scales in their establishing shots.

But first, I should note that shot scales have generally changed in movies over the last 70 years (Cutting & Iricinschi, in press; Cutting et al., 2012; Salt, 2006, 2009). In particular, the average shot scale has become shorter, in our data moving from roughly a medium long shot (~3.3) closer to a medium shot (~3.8). For these reasons, I normalized the shot-scale data within each movie then aggregated them across the 24 movies.

Several effects for shot scale are embedded within the data of the rightmost column of Table 1. Notice first that the mean non-normalized shot scale for null shifts (000) is considerably shorter than for any of the non-null shifts. This effect was a focus of Cutting et al. (2012) and of Cutting and Iricinschi (in press); shots within a scene are generally much shorter than establishing shots. Second and more important in this context, after shifts in location but not characters (101 and 100) the establishing shot has a longer scale than that after shifts of both location and characters (111 and 110) (~3.3) to (~3.8). For these reasons, I normalized the shot-scale data within each movie then aggregated them across the 24 movies.

Now reconsider the mean values in establishing shot scale when the environment changes but not the location (011 and 010) the establishing shots of the new scene have scales that are shorter than the average of those in which both shift (111 and 110), 2.69 vs. 2.15 (~3.3) to (~3.8). As noted earlier, this too makes sense. Given that the environment is unchanged the cinematographer need not step back and show the location, and given that a new character has entered it is prudent to offer the viewer a closer look at her or his face.

On the other hand, after shifts in which the character changes but not the location (011 and 010) the establishing shots of the new scene have scales that are greater than the average of those in which both shift (111 and 110), 2.69 vs. 2.15 (~3.3) to (~3.8). As noted earlier, this too makes sense. Given that the environment is unchanged the cinematographer need not step back and show the location, and given that a new character has entered it is prudent to offer the viewer a closer look at her or his face.

6. How stories are told in Hollywood style

6.1. Distributions of narrative shift types

How are these various shift types used in visual narratives? As a first step, I next assessed the percentages of these seven different types of narrative shift within each movie, and then averaged them across the

![Fig. 2. The segmentation concordance of observers across the 24 movies (3 observers per movie), under conditions in which the scenes are separated by cuts or by dissolves, fades, or wipes. The three integer places denote presence (1) or absence (0) of a location shift, a character shift, and a time shift, respectively, as denoted in Table 1. Error bars are one standard error of the means.](image)
24 movies. Results are shown in the third column of Table 1. Notice that there are striking differences across the set, well beyond what might be expected by chance, assuming that each shift type would have occurred one seventh of the time ($\chi^2(6) = 2764, p < .0001$). By far the most frequent narrative shift is that where both characters and location shift but time frame does not (110); 41% of all narrative shifts are of this type. The next most common is a shift where everything changes — location, characters, and time (111, 28%). Thus, locations and characters shift together across almost 70% of all narrative shifts in popular movies, a number that is fairly uniform in this sample over 70 years.

The third most common shift involves characters but neither location nor time (010, 13%), although as I discuss below this value has not been stable over release years. The fourth is its reciprocal involving only the location and time shifts but not characters (101, 9%), and the fifth concerns only location shifts (100, 7%). The final two, as noted above, are quite rare — time shifts alone (001, 1%), and shifts where both characters and time shift but location does not (011, <1%) — just as Messaris (1994) might have predicted. Coupled with their frequent and continued use of dissolves, their rarity adds force to the notion that these two shifts are difficult to understand.

6.2. A historical change in Hollywood style

Has the use of any of these narrative shifts changed over time? There is only one case — where only characters change (010) — and it is shown in Fig. 3 to have a sharp decline across release years ($r = .64, t(22) = 3.91, p = .0008, d = 1.67$). One major reason for this trend is that the older comedies and dramas were often adapted from live theater and retained their general economy of sets. Thus, characters walk in and out of a single location and without breaks in time. As they converse, the tone or the direction of the plot often turns. Notice also that this type of narrative shift is almost nonexistent in contemporary movies.

Another reason for the decline may be related to the increasing structural compactness of contemporary movies, called intensified continuity (Bordwell, 2006). At a shift in which only the characters change within a given location and time frame, something must account for the switch otherwise the change in social situation may appear rude or too coincidental without additional material. My guess, then, is that contemporary screenwriters and filmmakers may avoid such situations on grounds of streamlining the narrative.

6.3. Larger narrative effects on shift types

Movies are different in different ways. How do the patterns of narrative shifts change under the constraints of genre or through idiosyncrasies of the story line? Consider first genre differences. For those boundaries in which locations and characters both shift (111 and 110) there is a strong difference across the three genres ($\chi^2(2) = 440.2, p < .0001$). Table 1 shows that shifts of location and characters but not of time (110) are more common than those across all three dimensions (111). However, for the eight dramas the reverse is true, as shown in Fig. 4: Narrative shifts across all dimensions are more common. Given that this effect is uniform across the 70 years of filmmaking explored in this sample, I found this result surprising and worthy of pursuit.

To find probable cause, I estimated the lapse of diegetic time across each of the 24 movies, and enlisted several graduate students who were familiar with the movies to do the same independently of my estimates. I then log transformed the mean days estimated. Our results suggest that dramas entail a longer-duration story (detransformed mean = −260 days) than do comedies and action movies together (mean = −13 days, $t(22) = 3.03, p < .007$). For example, Goodfellas occurs across 15 years and Erin Brockovich across three years (both dramas); but Home Alone (1990) covers perhaps three days and Ocean’s Eleven (1960) probably a week (both comedies); while Beneath the Planet of Apes (1970) is likely only a few days, and Die Hard 2 (1990) perhaps only 8 h (both action movies). Thus, it seems sensible that because dramas unfold over a greater interval, more narrative shifts will incorporate shifts in time.

Consider next individual movies and narrative effects on shift type. Since Die Hard 2 (1990) takes place over only a few hours, one would expect few narrative shifts to invoke a temporal component (111, 101, 011, or 001) — and indeed, by my count, there are none. In contrast Goodfellas, as it unfolds over years, has many changes involving time, 125 of the 173 shifts (72%). Mission: Impossible II involves a typical protagonist/antagonist opposition found in action movies and one expects many subscenes interleaved with one another, changing characters and locations but not time (110) before a final climax. And indeed, 138 of the 184 shifts (75%) do just this. And in Harvey (1950), a studio-era adaptation from live theater, one expects — as noted earlier — that locations would stay relatively constant and have characters move in and out (010), broaching different issues as the plot develops; 36 of its 71 narrative shifts (51%) are of this type.

But one of the more interesting effects, as noted earlier, concerns movies with “reality” shifts (Messaris, 1994). Inception and Ordinary People are two of the three movies in this sample that shift in this...
way, and both do so extensively. Both movies use color quite successfully to denote these changes.9 Many diegetic scenes in *Ordinary People* show Conrad (Timothy Hutton) anguish over the death of his brother in a joint sailing accident. Intercut with these subscenes are visualized memories of that event, which are strongly tinted blue. In *Inception* the four different dream levels take place in a grayish, rainy urban landscape; a hotel interior paneled in lush browns; a mountainous landscape covered with deep snow; and the generally brown interiors of an apartment. Cutting et al. (2012) found that in both movies color change predicted viewer segmentations better than any other variable — shot duration, shot scale, motion, or luminance. *Inception* is particularly daring in its structure, with shifts from diegetic reality to and across dream levels and with the protagonist Cobb (Leonardo diCaprio) having memories that interpolate at two of these levels. By my counts the narrative traverses dream levels or goes from one dream level to a visualized memory 176 times across the course of the movie. Thus, one might wonder what aids the filmmakers have offered, other than color changes, to help viewers follow the narrative. The pattern of frequencies for the seven types of narrative shift in *Inception* is not markedly different than that for any other movie, with one exception. Instead of accounting for only 96% of all shifts, the shift of location and time but not character (101) accounts for 30% of them.

Only one other movie in this sample, *Grapes of Wrath* (1940), has a higher proportion of 101 shifts. That movie is the story of an extended family traveling together from Oklahoma to and through California in search of work. It seems inevitable, then, that locations would change with and over time, but that the characters would remain largely the same. Indeed 56% of all scene changes were of the 101 type, the most common type in that movie. Yet because *Grapes of Wrath* is an older movie with many fewer scenes (and no subscenes), it has only 25 of these changes, only about one quarter of the number in *Inception*.

In *Inception* there is also a scale difference in the establishing shot after such shifts. The scale after a 101 shift is nearly a medium shot (3.7), but the establishing shot after the other six shifts is closer to a medium-long shot (3.27, t(117.4) = 2.87, p < .005, d = .53). Moreover, in *Inception* these post-101-shift shots are tighter in scale than those in the other 23 movies of the sample (3.7 vs. 3.3, t(109.5) = 3.1, p < .0012, d = .59). Why might this be the case?

Of the 120 shifts across dream levels and memories in the last third of the movie, 58 (47%) are of type 101. Typically, the last shot of one subscene is fairly tight on a particular character at one dream level, and the establishing shot of the next is fairly tight on the same character at a different level. Given that viewers must assume that characters cannot be simultaneously in two places at once within the same dream or memory, these character-constant shifts (101) are strong visual signals that a dream-state or dream-to-memory state change has occurred. The use of these 101 shifts may be particularly important for those between dreams and memories — subscenes of which are often only one shot long. Indeed, 15 of 22 shifts (68%) from a dream level to one of Cobb’s memories and back are of the 101 type.

### 6.4. The movies of this sample and The Red Balloon

How do these results on the distribution of narrative shifts compare with those in the literature? The movie that has received perhaps the most psychological attention is *The Red Balloon* (1956), a charming 32-min movie with almost no dialog, about a boy and a balloon that follows him through a neighborhood of Paris (see Zacks & Magliano, 2011; Zacks et al., 2009; Zacks et al., 2010). Magliano and Zacks (2011) segmented the movie’s shots into three categories — those following continuity edits (n = 111), those following spatial–temporal discontinuities (n = 81), and those following action discontinuities (n = 19).

For this project I also segmented *The Red Balloon*, but according to the scheme used in this article and without first consulting Magliano and Zacks’s results. Although our categories differ, the correspondence of my scheme converted to theirs is almost perfect (r = .999, t(1) = 24.4, p < .026). The continuity edits correspond to the null-shift type (000; n = 102) plus the two least common shift types (011 and 010). Given that this movie has only two main characters (the boy and the balloon) my coding of the 011 and 010 shifts (ns = 1 and 6) involved those few shifts when the boy and the balloon are separated, wondering where each other is, but in essentially the same location. These can be regarded as consistent with continuity editing. The spatial–temporal discontinuities of Magliano and Zacks correspond to those shifts in which both the location and the characters change (111 and 110, ns = 6 and 13), regardless of whether there is a shift in time.

The percentage frequencies of the various shifts in *The Red Balloon* are shown in the fourth column of Table 1. Notice that they are considerably different than those aggregated across the 24 movies studied here, shown in the third column. Most of this difference may be accounted for by the fact that there are only two main characters in *The Red Balloon* versus the mean of ten in the sample studied here. Given that the boy and the balloon are almost always seen together there are very few 111 and 110 narrative shifts, which are the two most common in the movies of this sample. These shift types are also those that had the most reliable behavioral and fMRI results as reported in Magliano and Zacks (2011). Their former results are generally consistent with the results shown in Fig. 2.

### 6.5. Narrative shifts in movies and texts

Finally, how do narrative shifts in movies compare with those in text? Considerable work in event indexing and related approaches to discourse processing has looked at the effects on readers of shifts in locations, characters, time, and other factors in texts of various kinds (for example, Zacks et al., 2009; Zwaan, Langston, & Graesser, 1995; Zwaan, Magliano, & Graesser, 1995; Zwaan, Radavskys, Hilliard, & Curiel, 1998). In general readers are slowed by such changes as they update their mental representations of the ongoing narrative. Some analogous results appear to occur in movies as well (Cutting & Iricinschi, in press) in that establishing shots of a scene are longer in duration and scale if they present new information.

It was important in some of these works on text to note that shifts in the three indices studied here — location, characters, and time — are generally uncorrelated. This lack of correlation facilitated conclusions that shifts are additive in their effects on reading time (Therriault, Rinck, & Zwaan, 2006; Zwaan, Langston, & Graesser, 1995). Table 2 lists the correlations between these indices for these various works, which mostly used short stories as their stimulus material. In general these correlations were assessed within the units of a sentence or occasionally a clause.

How does one compare the textual results with those of movies? Analogies between text and film can be fraught, but there are few movie units that are contenders for the analog of the sentence in text. Film theorists (Metz, 1974), practitioners (Monaco, 1977) and psychologists (Carroll & Bever, 1976) agree that the shot is most like the sentence. Both are structural units, have discrete beginnings and ends (at

---

9. *Nine to Five* is the third film in this sample that has change-of-reality shifts of the manner that Messaris discussed. The three main characters Judy (Jane Fonda), Violet (Lily Tomlin), and Doralee (Dolly Parton) bond over their mistreatment by their boss, Frank (Dabney Coleman). In an apartment each day dreams of retribution and these are visualized and separated from diegetic reality by special wipes (iris ins and outs) and special effects — one dream is dark and in essentially black and white, another has animated cartoon creatures in it, and a third is staged as a western. There is no way one could mistake these for diegetic reality.

10. Most (32 of 49, 65%) of the location-only shifts (100) occur consecutively in the “gathering of the balloons” sequence just before the final scene of the movie.
Table 2
Correlations among location, character, and time shifts in text and in movies.

<table>
<thead>
<tr>
<th>Source</th>
<th>Location/characters</th>
<th>Location/time</th>
<th>Characters/time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Text</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zwaan, Langston, and Graesser (1995)'</td>
<td>.16</td>
<td>.29</td>
<td>.13</td>
</tr>
<tr>
<td>Zwaan, Magliano, and Graesser (1995)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experiment 1'</td>
<td>.15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experiment 2'</td>
<td>.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zwaan et al. (1998)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experiment 1</td>
<td>.22</td>
<td>.35</td>
<td>.28</td>
</tr>
<tr>
<td>Experiments 2 &amp; 3'E</td>
<td>−.03</td>
<td>.32</td>
<td>.30</td>
</tr>
<tr>
<td>Zacks et al. (2009)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experiments 1, 3, and 4'</td>
<td>.15</td>
<td>−.05</td>
<td>.00</td>
</tr>
<tr>
<td>Movie</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zacks et al. (2009)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experiment 2g</td>
<td>.37</td>
<td>.26</td>
<td>.21</td>
</tr>
<tr>
<td>The 24 films in this sample'h</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean &amp; 95% confidence interval</td>
<td>−.06 ± .05</td>
<td>.16 ± .07</td>
<td>−.09 ± .07</td>
</tr>
</tbody>
</table>

*a Unit = the sentence in short narratives.
*b Unit = the sentence in two short stories.
*c Unit = the sentence or clause in two different short stories.
*d Unit = the clause in adaptations of some of Aesop’s fables.
*e Unit = the clause in narratives used by Rinck and Bower (1995).
*f Unit = the phrase in narratives taken from a descriptive diary.
*g Unit = consecutive 5 sec intervals in The Red Balloon (1956).
'h Unit = the shot in 24 popular movies.

least with cuts), can be of almost any length, and can carry prodigiously different amounts of information. Moreover, for decades the average shot in the average movie took the same amount of time as it takes an average reader to read an average sentence. That is, if an individual reads at 200 words per minute and if the average sentence is about 15 words (Cutts, 2009), then the average sentence will be read in 4.5 s. This is the duration of the average shot for movies released between about 1975 and 1995 (Cutting, Brunick, Delong, Iricinschi, & Candan, 2011).

The methods and content used across these other studies and this one are quite varied so it is difficult to compare results. For example, few of the texts used in the discourse literature have dialog, one character speaking after and in response to another. If they did, these might be coded as having a character change after every utterance. In popular movies, on the other hand, a large proportion of scene time is devoted to dialogue, which is typically done in shot/reverse-shot (reverse angle) style. That is, often after an establishing shot showing both speakers, the camera alternates focus on one person, then on the other. I have not counted these as character changes, and because of expectations in turn-taking (e.g. Sacks, 2000) viewers would be surprised if the characters did not alternate in what they say.

In addition, there are differences according to how the correlations are calculated. If the correlations are calculated on every shot change (or every word change) the high probably of no shift after every shot (or word) inflates the correlations. If however the correlations are done on every event boundary, the correlations decline dramatically. Since the correlations in the text domain were done on event boundaries, I report those for movies as well. With these caveats in mind and as shown in Table 2, the correlations among shifts in location, characters, and time in the sample of 24 movies studied here are reasonably in line with those of the various texts from short stories and diary accounts used elsewhere in the discourse literature.

7. Summary and conclusions

Seven types of narrative shifts were described involving the presence (1) or absence (0) of changes across locations, characters, and time, respectively — 111, 110, 101, 100, 011, 010, and 001. Several patterns were decanted from observers’ segmentations aligned with these types of narrative shifts in 24 movies. First, the more dimensions — location, characters, or time — that change, the more consistently viewers segment. Second, although observers are somewhat more consistent when the changes are accompanied by dissolves, fades, and wipes, they are also quite consistent without them. It seems likely that the marginal loss in segmentation performance would be outweighed in the minds of contemporary filmmakers by the importance of accelerating the narrative with fewer noncut transitions. Moreover, I suggest that a certain amount of perceptual elision between scenes actually contributes to continuity.

The relative frequencies of the seven shifts are quite varied. Location and character shifts have a strong tendency to co-occur (111 and 110) and these types are by far the most common, accounting for about 70% of all narrative shifts. Time changes are somewhat more independent of locations and characters, and their frequency of occurrence seems to depend more on the duration of diegetic time within the movie; the longer the period the movie covers, the more often they appear. In particular, dramas tend to have more shifts across time than do either comedies or action movies. There are, however, idiosyncratic patterns of shifts in many movies that derive from special aspects of their narratives. The Red Balloon, intensively investigated by Zacks and colleagues, is one such case.

Some types of narrative shift are quite rare. The least common are those where location does not change, time does change, and characters either change (011) or do not (001). Together these account for less than 2% of all shifts in this sample. Moreover, filmmakers seem to know that they are difficult to process since about two-thirds of those that do occur are accompanied by a dissolve, the most traditional method of signaling a time change. Dissolves are otherwise rare in contemporary movies, and narrative shifts in which the characters change but the location and time do not (010) have become rare as well, marking a difference in Hollywood style from 60 and 70 years ago.

Finally, shifts involving location changes without character changes (101 and 100) tend to be followed by an establishing shot with longer scale (more of the environment is shown) than others, and changes involving character shifts alone (011 and 010) tend to be followed by an establishing shot with a shorter scale (offering more of a close up of the new character). The intermediate shot scale values for changes where locations and characters (111 and 110) shift together seems to result from a compromise between moving out to show the new environment and moving in to show the face of the new character.

Together these results suggest that there are four different signatures of narrative shifts to be found in popular movies — general patterns across time, historically changing patterns, genre-specific patterns, and film-specific patterns.

In addition these results also support our previous claims (Cutting, Brunick, & Delong, 2011; Cutting & Iricinschi, in press; Cutting et al., 2010; Cutting et al., 2012) that the structure of popular movies, and the changes in that structure over the last 80 years, are fit to perceptual and cognitive processes that allow movies to be faster paced but still easy to understand. Moreover, we know that in real life and in movies this segmentation process is an aid to memory and comprehension (Zacks et al., 2009; Zacks et al., 2010). Thus, understanding the reasons for why and how viewers segment movies into events brings us closer to understanding why we find them such a powerful component of contemporary visual culture.

Acknowledgments

Thanks to Kate Brunick, Ayse Candan, Jordan Delong, and particularly Catalina Iricinschi for earlier work, discussion, and guidance assembling all of this material; to Michael Masucci for his early help in scene coding; and to David Field, Khena Swallow, James Golden, and Kedar Vilankar for broader discussion.