DE GRUYTER MOUTON
DOI 10.1515/sem-2012-0091 — Semiotica 2012; 192: 123–160

Chiao-i Tseng

Audiovisual texture in scene transition

Abstract: This paper shows how the method of cohesion analysis in film provides a strong foundation for constructing coherent filmic narrative across different spatial and temporal settings. Until now the unit of a fine-grained systematic analysis of film viewers' narrative comprehension has been anchored in the shot. This paper provides a semiotic descriptive tool of filmic cohesion for analyzing elements within and across shots, i.e., how viewers comprehend the presenting and tracking of dominant narrative elements such as characters, objects, and settings as a film unfolds. In addition, this paper also reviews the Neo-Formalist exploration of cinematic cohesion and then demonstrates how the semiotics-oriented method suggested here can support more systematically and even strengthen the Neo-Formalist account. Most importantly, this article will explicitly show the potential of this tool for providing empirical accounts of filmic narrative construction by analyzing scene transitions in Wong Kar Wai's film *2046* (2004).

Keywords: film analysis; multimodality; discourse semantics; cohesion; language and film; texture

Chiaoi Tseng: Bremen University. E-mail: chiaoi@uni-bremen.de

1 Introduction

Scene transitions in film indicate the manipulations of changes from one setting to a different setting, or from one time frame to a new time frame. In the historical development of meaning analysis in film, how viewers are carried from scene to scene has been a frequently addressed question. The study of the manipulations and effects of scene transitions marked one significant turning point – for some, investigating the construction across different scenes and shots provides insight into how meanings in film are created and signaled to the viewers. Eisenstein was one of the first influential theorists in this regard. Eisenstein views the combination of different types of shots and scenes as a kind of hieroglyph or pictogram: that is, the filmmaker uses the editing process to combine shots or scenes of different spaces and times in such a way that an abstract idea is conveyed to the

audience by means of metaphor and the association of ideas (Eisenstein 1969). In the 1960s and 1970s, transitions of shots and scenes were probed in a more scientific fashion by the film semiotician Christian Metz. Metz systematized the constrained ways of ordering narrative space and time between shots and scenes and developed his well-known model of the *Grande Syntagmatique* (Metz 1974: 146). He divided the narrative syntax of the cinema into eight structural configurations, whereby the viewers make sense of cinematic structures. Following the pursuit of Metz to create a generalized modeling of filmic meaning, the recent study of Bateman (2007) reviews the historical impacts and debates building on the Metzian semiology since the 1970s and revises Metz's *Grande Syntagmatique* into the *Grande Paradigmatique*, which maps out a more comprehensive set of resources for manipulating semantic relations such as different types of temporal, spatial, logical relations; these resources account for just how the viewers are cued to the coherent narrative flow across different shots and scenes.

These developments have hitherto shown that the theorizing of the mechanisms for cinematic transitions across different spatial/temporal frames provides a significant analytical basis for the empirical inquiry into the construction of filmic meanings. Drawing on its theoretical value in this respect, this paper aims to do three things. First, apart from the mechanism of shot relations investigated in the past eighty years, this paper addresses the question of whether we can also probe into another dimension of mechanisms, namely, the resources for establishing the audiovisual texture when the transitions of spatial or temporal frames take place. More specifically, it addresses how elements within film images, whether realized in the audio- or visual-track, are tied together to signal the coherent flow of film narratives across different shots and scenes. In this pursuit, this paper first unravels the significance of the texture dimension in terms of viewers' comprehension of filmic meanings and suggests an analytical method for describing the audiovisual texture in moving images.

Second, this paper reviews and strengthens the exploratory attempt made by David Bordwell (2008) in a similar direction. In his essay "The hook: Scene transitions in classical cinema," he provides an empirical account of how audiovisual cohesion is highlighted and how patterns of film cohesion unravel viewers' cognitive activities as a film unfolds. Bordwell generally terms the devices for managing scene transition in this dimension "the hook." Building on his cognitive model and poetics of cinema (Bordwell 1985, 2007), Bordwell reviews the history and functions of the hook and examines some possible options and structures for realizing different types of the hook across scenes. This paper treats Bordwell's attempt to identify internal cohesive structure in film as its point of departure and shows a more systematic analysis of audiovisual cohesion that the hook yields.

Finally, and most importantly, the tool for analyzing audiovisual texture is applied to examining the scene transitions in Wong Kar Wai's film *2046* (2004). The story centers around a man, Chow, recalling his past relationships with women in Hong Kong and Singapore in the 1960s. It contains at least three main story strands not only from 1963 to 1970 but also in Chow's imagination. What is particularly intriguing in *2046* is the non-chronological intersection of these story strands. The non-linear narrative unfolds in a complex fashion, which echoes the protagonist's wild imaginations about a science fiction story he is writing and his memory of the past. Despite the non-linearity of the narrative, this cohesion analysis will show just how the cohesive devices are mobilized to set the viewers on a certain path of narrative interpretation. This analysis of *2046* thus will demonstrate more concretely the potentials of this tool for providing empirical accounts.

2 The notion of texture

The last section began with a brief overview of shot-based approaches to filmic meaning creation. From Eisenstein and Metz to Bateman, the unit of analysis has been based on narrative meaning constructed by a shot. The following sections provide insight into another dimension of meaning construction, namely, *texture*. The notion of texture applied in this paper has a linguistic origin. When it is applied to film analysis, it is anchored in the cohesive devices for cueing viewers' comprehension of elements *within and across* a shot (cf. Palmer 1989: 316). This dimension addresses, in particular, how characters, objects, and settings, whether realized in the visualtrack (e.g., visible figures or as written names on the screen) or in the audiotrack (e.g., spoken names or sounds and music that represent certain identities), are coherently signaled to the viewers throughout a film.

The significance of such analytical units in terms of meaning construction and comprehension has been articulated by the cognitive film theorists. Bordwell, for instance, explains the cognitive schema termed "bull's-eye schema" (1989: 171), which puts characters, characters' actions, and the relationships between characters as the most important narrative interpretive cues. Less salient but still significant are the characters' settings and their relationships with objects in the surroundings they inhabit (Bordwell 1989: 170). Building on the interpretive significance as such, the aspect of texture to be focused on in this study is precisely how the coherence of these interpretively significant elements such as characters, objects, and settings, is signaled to the viewers; and furthermore,

how such texture analysis can be used empirically, for instance, in addressing stylistic issues of scene transition.

2.1 Texture in language

Generally, in linguistic analysis, texture refers to the mechanism that makes any length of text meaningful and coherent. Eggins (1994: 95), for instance, uses the linguistic concept "sequential implicativeness" put forth by Schegloff and Sacks (1973: 74) to explicate that language follows a linear sequence where one line of text follows another with each line being linked or related to the previous line. This linear progression of text creates a context for meaning; texture, in this respect, is the basis for unity and for the construction of a context of meaning. A text without texture would just be random bits and pieces of isolated words or sentences with no relationship to one another.

Texture is realized by the deployment of cohesive devices, i.e., the semantic ties within text whereby the interdependency of textual elements is created. As a major step in the development of Systemic Functional Linguistics, Halliday and Hasan (1976) sketch five types of cohesive mechanisms which create texture in text – reference, substitution, ellipsis, conjunction, and lexical cohesion. Since the dimension of texture to be addressed by this paper is how a previously mentioned/shown element in film, whether a character, an object or a setting, is coherently referred to again across different shots and scenes, the linguistic notion most suitable for application is the framework of "cohesive reference," also reformulated as "identification" by Martin (1992) as a system of discourse semantics of language.

According to Martin (1992: ch. 3), the choices of the identification system in natural language refer to devices of identification functions to retrieve presupposed information in text. These devices indicate how the text producer introduces people, places, and things and keeps track of them throughout the text. Whereas Martin sees this resource as a component of discourse semantics for language, my previous work (Tseng 2008, 2009) takes this further and develops an identification system specifically for analyzing moving images. The application of functional linguistics to film analysis not only provides a powerful tool for examining cohesive ties between film elements within and across film image frames but also strengthens the film-language analogy, a theoretical correlation that has been pursued for the past eighty years. More importantly, the systems formulated specifically for film discourse (rather than directly equating language and film systems) show that it is nevertheless necessary to differentiate between the film and the language systems.

2.2 Analyzing multimodal texture in film

The following account examines the filmic identification system, which describes the mechanisms that cue the viewers to the presentation and reappearance of people, places, and things in film, i.e., characters, settings, and objects. The perennial debate about how viewers identify with characters in film, and about their levels of emotional engagement, falls outside the scope of this article (on methods see, e.g., Smith 1995; Eder 2008, 2010). Instead this article concentrates on the sorts of analytical systems that allow propositions to be formulated about the viewer's comprehension of characters, objects, and settings as continuous and re-identified (referred to as identity presuming in this article), as either individuated or generalized, and as a driving force for constructing coherent narratives.

The filmic identification system is presented in Figure 1, modeled in a paradigmatic system. In the theory of systemic functional linguistics, such system networks are used to show the abstract paradigmatic "choices" available for

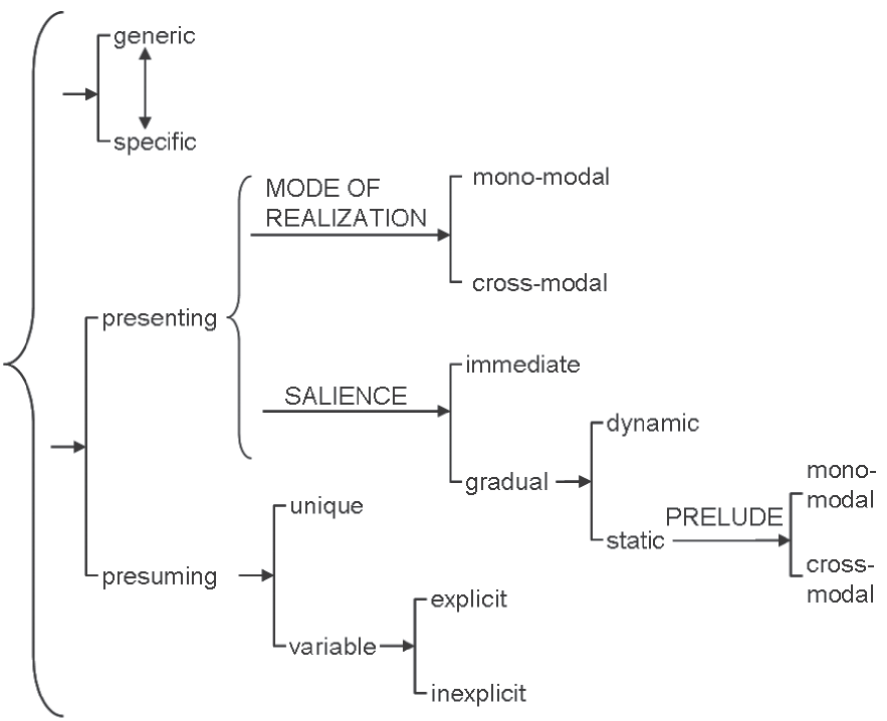


Fig. 1: System network of filmic cohesive reference

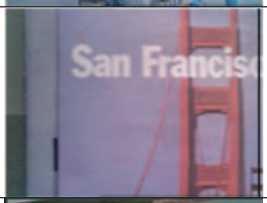
1		**ShotA**. (cable car bell ringing + squawking)
2		(Car honks + squawking)
3		(squawking)
4		(car honks + squawking)
5		*San Francisco*
6		(squawking)
7		(wolf whistle)

Fig. 2: Transcription of the first fourteen images across two shots in The Birds. (Transcription conventions: ( ) = sound effect, italic = written text, " ": spoken text)

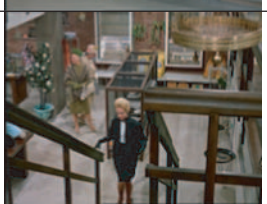
8		(squawking)
9		(squawking)
10		(gull squawking)
11		(gull squawking)
12		*Davidson's Petshop*
13		**ShotB**.(chirping)
14		(chirping) Melanie: "Hello, Mrs McGruder". Shopkeeper: "Hello, Mrs Daniels". Melanie: "Have you ever seen so many gulls?"

Fig. 2: (Cont.)

language users drawn from the meaning potential of their language (Halliday and Matthiessen 2004).

The filmic system in Figure 1 shows the functional potential of cueing the same character, object and setting throughout a film, whether within or across a shot or a scene. In these networks, square brackets connect contrasting options together into systems: for instance, in the system of [presenting (i.e., choices for introducing an identity for the first time)/presuming (i.e., resources for tracking a previously presented identity)] only one of the two features may be selected at a time. The system networks can also employ simultaneously available systems represented by grouping systems together with a curly right-facing bracket. In Figure 1, for example, choices need to be made from the features presented by both the systems [generic/specific] and [presenting/presuming]. That is to say, one of the following four types of combination must be selected: (1) [presenting] + [generic]: e.g., introducing a general group of children for the first time, (2) [presenting] + [specific]: e.g., introducing a specific character for the first time, (3) [presuming] + [generic]: e.g., re-identifying a general group of children which appeared previously and (4) [presuming] + [specific]: e.g., re-identifying a main character. Analogously, one option from each of the systems of MODE OF REALIZATION and SALIENCE needs to be selected, thereby giving rise to sometimes quite extensive cross-classification.

To illustrate the operation of these networks for cohesion in films, I will first exemplify the presentation and re-identification of the main character, Melanie, in the first fourteen images of the Hitchcock film *The Birds* transcribed in Figure 2.

In the beginning of this film, Melanie walks to the camera from the background (from image 2 to image 6) – in image 2 she first stands in a group of people at a distance from the camera, walks to the left side of the image where she becomes isolated and salient; and then she walks behind the *San Francisco* poster and finally appears as a foregrounded character in a medium shot. In image 11 she notices the squawking gulls before entering into a pet shop in image 12.

The cohesive devices instantiated from the system for presenting and tracking her identity can be described as follows – in image 2, this female character is presented for the first time and thus in this image it is appropriate to make the choice of the feature [presenting] from the system [presenting/presuming] to describe this; moreover, she is presented only visually (not simultaneously in written or spoken text), thus the [mono-modal] realization from the system [mono-/cross-modal] is also selected. The foregrounding of her appearance is a *gradual*, as she gradually moves from the background to the camera. Hence, this *gradual* process can be seen as realized by the choice of [gradual salience] rather than [immediate] in the system SALIENCE.

The network in the figure also shows that the choice of [gradual salience] leads on to finer classifications. This option classifies a presentation strategy whereby a character/object is progressively foregrounded. In the present sequence, where the character is placed in the background at the outset of the presentation and then is gradually foregrounded by the character's *dynamically* approaching the camera from a distance (from image 2 to image 6) – in this case, the choice of [dynamic gradual] is selected.

As Figure 1 also shows, in other cases the choice of gradual salience can be realized not with the characters moving towards the cameras, but by some other kind of hinting elements termed *prelude*, whether realized mono-modally or cross-modally, for signaling the upcoming appearance of characters. For instance, seeing some physical parts of characters/objects or hearing some sounds these characters/objects produce before their identities are fully and explicitly shown, viewers are cued to the possible existence of characters/objects. There is such a prelude in the present sequence realized in the presenting of "birds" – although the birds are visually shown in image 10, the audio prelude "squawking sounds" can be heard from image 1 to image 9; and these sounds hint the upcoming visual presentation of the birds.

Another choice must be made from the system of [generic/specific]: the two features available here refer to the degree of generality of identities of characters, objects, and settings. This system is modeled as a continuum rather than as contrasting options: realizing generic and specific identities in film is not an either/or choice, but a continuum of relative degrees of generality, varying from "the most general characters, objects, and settings" to "specific individual identities." This generality of identities can be manipulated in film by several strategies. For instance, a character wearing certain visual attributes that represent specific social types (e.g., the scarf worn by Muslim women) in a viewer's culture is regarded as less generic than any character without social cultural cues. Moreover, a generic character, although unnamed or unlabelled in the narrative, can be gradually "specified" when he or she repeatedly appears and is recognized by viewers as a certain specific character. Within the current segment, the female character is simply presented visually in images 2–3 and hence there are at first no grounds for assuming that she is anything more than generic. To sum up, the cohesive relation describing Melanie's initial presentation is then:

[generic identity] + [presenting] {[mono-modal realization] + [gradual, dynamic salience]}

These are the filmic devices of presenting her identity employed in the sequence. After she is visually introduced in image 2, she is then seen moving from

image 2 to image 6 gradually towards the foreground and reappears explicitly from image 7 to the end of the sequence. This is sufficient to move her identity from "a [generic] female character" to "the [specific] female character" as the film proceeds.

Descriptions of the identification strategies employed in a film can be constructed in a similar way for all narrative elements (i.e., characters, objects, settings) that might appear. The network in Figure 1 captures the range of possible strategies supported by film, while the actual features selected in any segment characterize the individual film in relation to the background offered by the potential as a whole. Carrying out this analysis for the remainder of the *The Bird* extract provides a detailed characterization of the ways in which the identity of each prominent character, object and setting is realized, presented, and presumed (tracked) drawing on this overall potential.

This form of analysis is then taken further by setting out the cohesive strategies adopted across a filmic segment for individual identified elements. This yields cohesive chains that are formed whenever particular elements are placed repeatedly in sequences of cohesive ties over the unfolding of the film sequence. Whereas any element in a textual artefact typically enters into a large number of cohesive links with other elements, it has been observed in work on language texts that a particularly strong textual role is played by cohesive chains rather than individual elements. Within the example extract being analyzed here, seven prominent narrative elements of character, object, and setting can be identified: *car, the San Francisco street view, Melanie, people on the street, birds, petshop, shopkeeper*. Each of these elements in a cohesive chain is made up of a sequence of cohesive relations. Other narrative elements that may have been relevant due to their in frame presence fall away at this point precisely because they do not participate in a chain (cf. Hasan 1984).

The seven cohesive chains obtained by collecting cohesive ties for this segment are shown in Figures 3, 4, 5, and 6. The chain pattern is broken into four diagrams only due to layout constraints, but Figure 6 shows the overall shape of the pattern formation.

Each of Figures 3 and 4 shows three strands that when taken together track the realization of each of the first three identities operating within the extract from image 1 to image 14: the object of "car" on the street, the setting of "street view," "Melanie"; while Figures 5 and 6 show the other four identities in this extract: general "people on the street," "birds," the second setting of "pet shop," and the "shop keeper."

The maintenance of each filmic identity chain is then shown using arrows that link successive elements back to previous elements of the same chain; and in these figures, each instantiated choice from the system network is particularly specified next to the link along each chain.

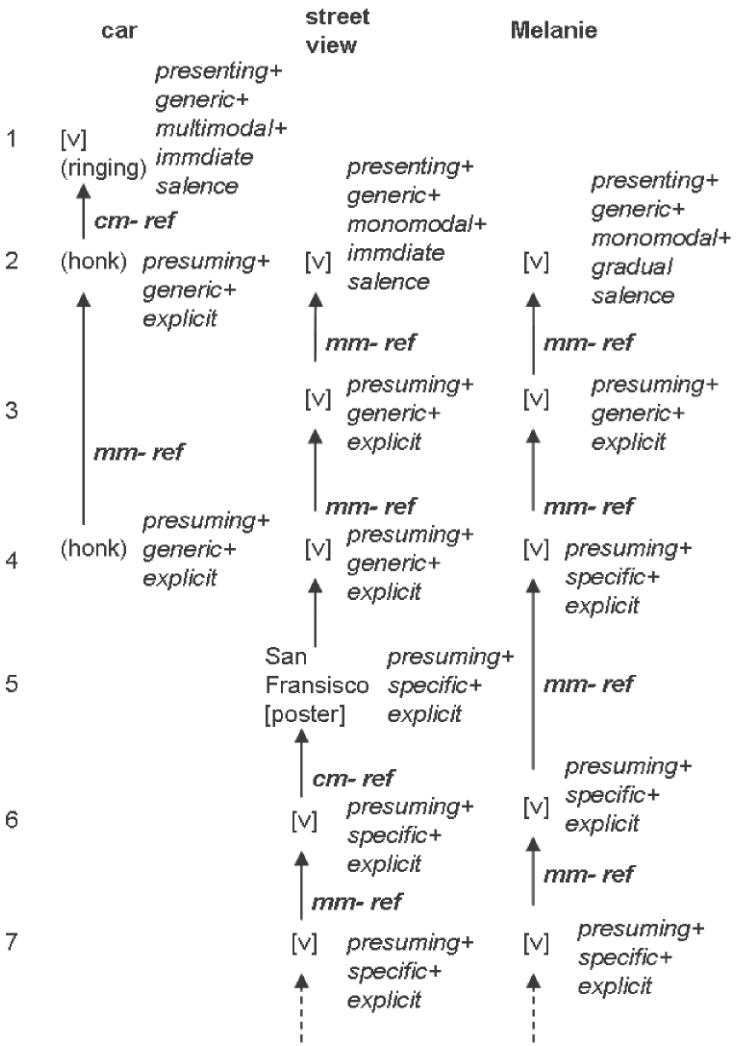


Fig. 3: Chain patterns images 1–7: identity chains of "cars," "street view," "Melanie." Transcription conventions: [v] = visually realized characters, objects and settings, ". . ." = spoken text, ( ) = sounds, numbers on the left give the shot number, italic text: the choices instantiated from the identification system, bold and italic text: cm-ref = cross-modal referencing, mm-ref = mono-modal referencing

Starting again with Melanie, the main character (which we will now be able to identify as the main character on the basis of her reoccurrence in cohesive chains), as discussed earlier, she is first presented in images 2–3 as a generic

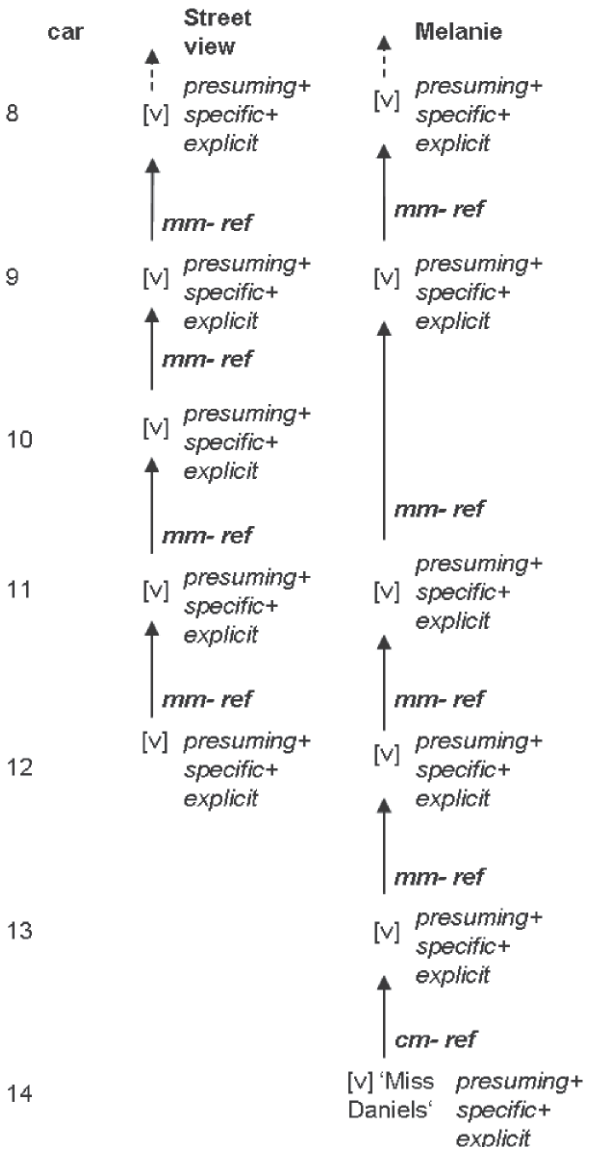


Fig. 4: Chain patterns images 8–14: continuing the identity chains of "cars," "street view," "Melanie." Transcription conventions, as Figure 3.

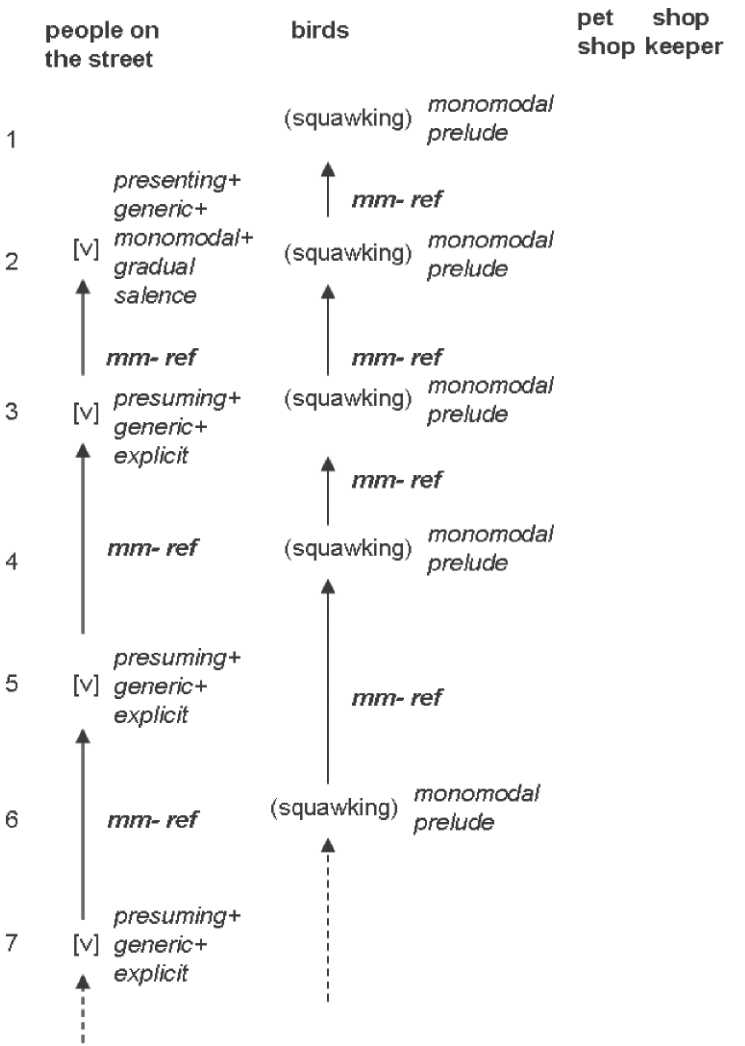


Fig. 5: Chain patterns images 1–7: identity chains of "people on the street," "birds," "petshop," "shop keeper." Transcription conventions, as Figure 3.

female character, monomodally in visual image, and her movements towards the camera gradually signal the salience of this person. From image 3 to image 14, her identity is tracked repetitively and explicitly in visual images; and in image 14, her identity is also named in the spoken text ("Miss Daniels") by the shopkeeper.

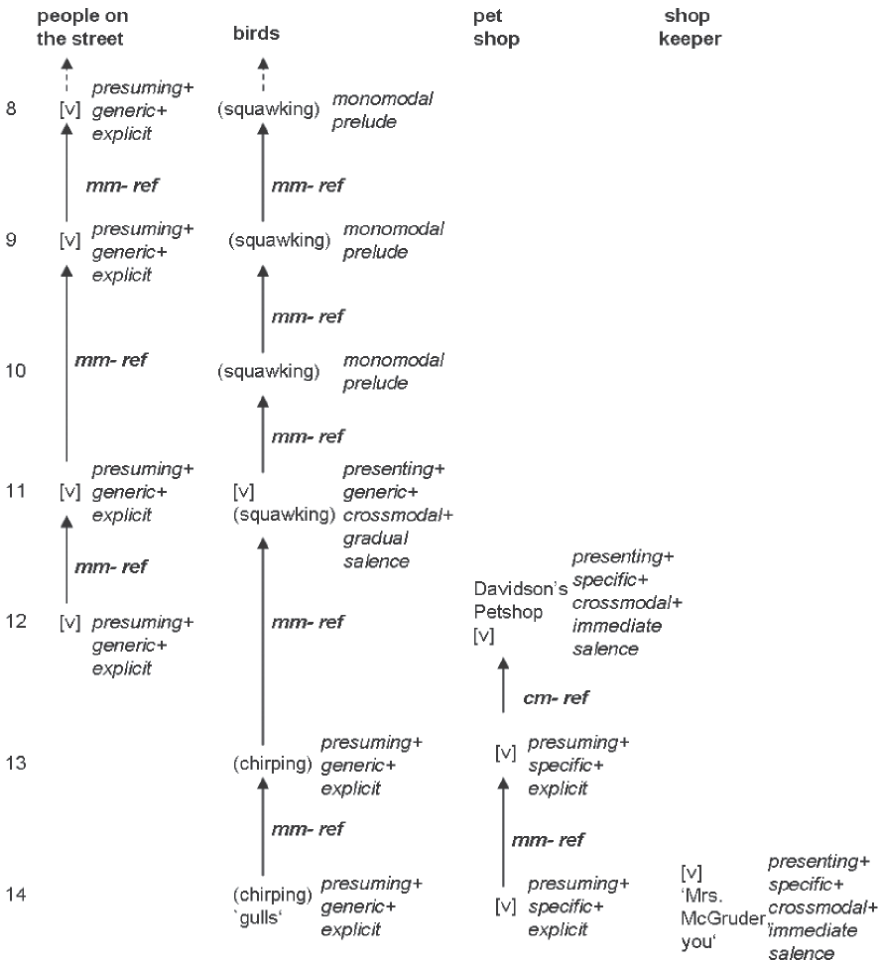


Fig. 6: Chain patterns images 8–14: continuing the identity chains of "people on the street," "birds," "petshop," "shop keeper." Transcription conventions, as Figure 3.

Hence, the two choices realizing the presuming (tracking) from image 3 to image 14 along the Melanie chain are: (1) [specific identity] from the system of [generic/specific] because the visual reappearance has converted her generic identity of "some female character" to "the female character," and (2) [explicit reappearance] under the [unique/(variable explicit/inexplicit)] system because her identity is *not* a [unique] identity that needs to be widely-known in the viewer's culture, and the reappearance of this female character's identity, whether verbally or visually, is realized *explicitly*.

Referring to these figures, we can also see that the tracking of some identities is realized cross-modally. For instance, the first chain generic "car" starts in image 1 with a cross-modal realization both in the ringing sound and the image of the cable car, thus specified as (ringing) and [v] in the chain; and in image 2, the identity of "car" is tracked again in the audio mode, as a honking sound this time. Hence, in Figure 3 the sign of **cm-ref** is added next to the arrow specifying the cross-modal referencing in the process of identification of this object. In image 4, the same honking sound can be heard again, referring back to the audio mode in image 2 through mm-ref, namely, mono-modal referencing, realized only in audio mode.

The second chain tracks the setting "San Francisco street view." It is presented in image 2 in the visual mode (thus "[v]" in the diagram) and then realized in the following images as generic setting until image 5, where the verbal element in the poster confirms its specific identity. Furthermore, through examining this chain and the next setting chain "pet shop," a clear setting transition is highlighted – the setting chain of San Francisco continues until image 12, as Figure 3 shows; and the second setting chain "Pet shop" starts right in image 12, as the chain displayed in Figure 6). The general formation of chain pattern in Figure 6 can highlight this transition of spatial-temporal context more explicitly, that is, the "pet shop" chain starts right after the "San Francisco street view" chain ends.

The third chain of this extract (shown in Figure 3 and 4) represents the presenting and tracking of the main character "Melanie." The diagrams display that the first element of the chain is realized by the choices of *monomodal* (visual), *generic presentation* with *gradual salience*, as discussed earlier in this section. The strategy of match-on-action used across images 12 and 13 in this extract signals a continuous movement of Melanie who enters into a pet shop from the street. With this continuity device, the Melanie chain remains unbroken throughout the extract despite the scene transition in image 12, as can also be seen in the overall pattern of Figure 7.

Figure 5 and Figure 6 display the construction of another four identity chains throughout this extract: "general people on the street," "birds," the second setting of "pet shop," and "the shop keeper." Here the choices specified next to each link of the chain "people on the street" shows that this character type is presented *monomodally* to viewers as *generic* throughout the extract – they are visually presented in the background and no specific characters among them are introduced to the viewers. The next chain links the presenting and tracking of another dominant role in this extract: "birds." As discussed earlier, the squawking sound functions as a *prelude*, heralding the upcoming birds, which are to be visually presented later in this extract. Hence in the figure, each link of the "bird"

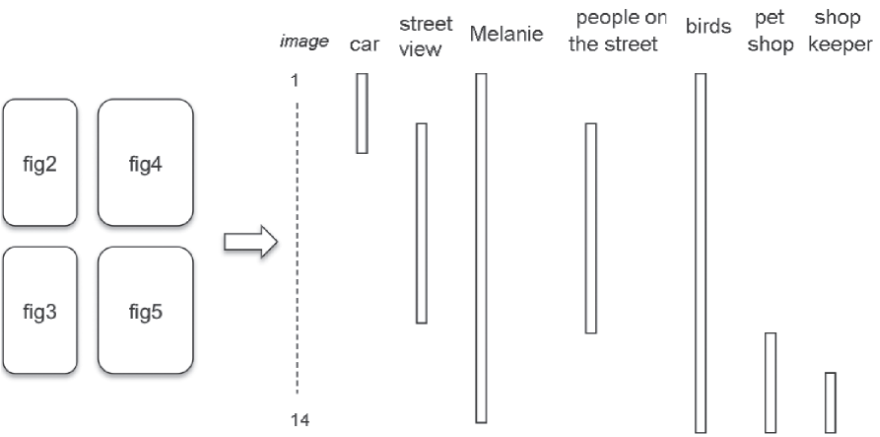


Fig. 7: The overall pattern of the cohesive chains in this extract mapped in a more abstract form.

chain from image 1 to image 10 is specified with its function role as prelude. Finally, the last two chains shown in Figure 6 comprise elements appearing only in the second scene of this extract, the setting of pet shop and the shop keeper. As Figure 6 displays, the "petshop" chain is presented simultaneously in visual and verbal modes (i.e., the written signs, "Davidson's Petshop") and then tracked in the following images in the visual mode. The last chain "shop keeper" is also presented multi-modally: visually as a female character and simultaneously called (by Melanie) "Mrs. MacGruder."

The overall formation of the chain pattern in Figure 7 shows the different degree of dominance of the characters. While the chains of Melanie and bird are dominantly constructed across the extract, the general group of people on the street appears only briefly in the middle of the extract. Moreover, that "street view" chain is followed by the construction of the "pet shop" chain, which also demonstrates more clearly how the scene transition takes place and which are the dominant character and object re-identified across the two settings despite the scene transition.

To this point, I have explicated in detail how to analyze the cohesive devices mobilized as a film unfolds. The chain patterns displayed in Figures 3, 4, 5, and 6 show how the identity of each dominant character, object, and setting of this extract is presented to the viewers and is realized throughout the extract to construct a coherent film narrative. In order to explicitly unravel just how the features in the paradigmatic system of filmic cohesive reference (Figure 1) are instantiated, along each cohesive chain I also specified the features selected from the system next to each link of the chain.

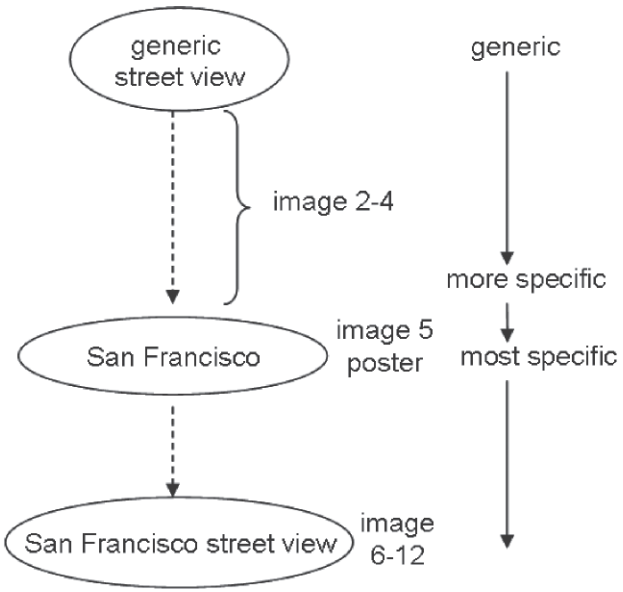


Fig. 8: Dynamism in viewers' interpretation process (arrows indicate the direction of the gradual revision)

More importantly, chain patterns as such reveal the dynamic potential for showing viewers' ongoing revision of meaning inference. Such dynamism is articulated in Figure 8, taking the viewers' comprehension of the San Francisco street view for an example. The semi-circle shows the viewers' interpretation as the extract unfolds. It begins with the most generic street setting, and the use of the device "visual repetition" moves the identity of the setting from the end of [generic] gradually to the other end of the continuum [specific]; and finally, the insertion of the poster in image 5 further confirms the identity in the most specific way. From image 5 onwards, viewers' interpretation of the same street views is then specifically directed to the identity of San Francisco. This diagram thus shows how the tracking of an identity drawing on cohesive ties can actually untangle the dynamic aspect of meaning articulation as a film discourse unfolds and viewers' ongoing adjustment of narrative inferences.

The potential of showing the dynamic aspect as such further brings to light the need for reconceptualizing the problematic dichotomy of semiotic codes and inference. This dichotomy is anchored in the parallel typically drawn between semiotic codes and static, non-inferential syntax. Tseng and Bateman (2010) explicitly illustrate the fallacy of this conceptualization:

> Our reconstrual returns us to discourse and dynamism, and to the, in our opinion, incorrect assumption that semiotic codes are only a matter of static rules and non-inference. In particular, we realign semiotic code with discourse and dynamism rather than with the inherently contradictory pairing of discourse and staticity. To do this, we build directly on more recent views of discourse organization and interpretation developed in both formal and functional linguistic work on discourse representation and discourse semantics. (Tseng and Bateman 2010: 216–217)

The present paper will explicitly explain how the cohesion analysis at the level of discourse semantics combines the semiotic codes and viewer's inferential activity. This point will be further elaborated in a later section where this method will be applied to analyze the viewer's inference construction of *2046*. But before that, the next section will give an overview of the account provided by David Bordwell on cinematic cohesion and scene transition drawing on his neo-formalist framework.

3 The hook in Neo-Formalism

The descriptive realm of Neo-Formalism, which this section reviews, is an approach developed by Bordwell and Thompson (Thompson 1988; Bordwell 1997, 2006, 2007). The methodological principle of this analytical framework is generally a top-down approach to schemata construction complemented by bottom-up investigation of the functioning filmic devices in a film. From this perspective, perceptual data, namely, narrative films, are conceived as a set of cues interacting with the spectators' cognitive capacity, triggering and constraining their activity of inference generation. In Bordwell's words, "[the cues in the narrative film are organized in such a way as] to encourage the spectator to execute story construction activities. The film presents cues, patterns, and gaps that shape the viewer's application of schemata and the testing of hypothesis" (Bordwell 1985: 33).

Thus, one main task of film analysts is to describe the cues and their functional roles in triggering and constraining viewers' comprehension. He reuses the term poetics of cinema to specifically denote this descriptive realm. Since the 1970s, Bordwell and Thompson (Thompson 1988; Bordwell 1997, 2006, 2007) have published a wide range of works describing the functions and stylistic conventions of filmic devices which cue spectators to a coherent filmic narrative for comprehension, as well as how these cues are used differently and recurrently across epochs.

When applying this principle to the empirical exploration of cinematic cohesion, Bordwell (2008) proposes that cohesion be taken as one of the three broad levels of structural devices that reflect film coherence:

- The first large-scale structure is macrostructure whereby acts in a film hang together. Each act represents a phase of certain goal-oriented story action. For instance, the traditional three-act macrostructure encompasses three story phases: presenting the main characters with certain goals, the main characters being blocked in achieving these goals, and finally, at the climax the main character decisively achieving the goals (or not).
- The mid-level structure is plot coherence, concerned with how scenes or sequences in film are connected, e.g., the cause-and-effect relation between these sequences. At this structural level, scenes are seen as aiming for advancing the action and for connecting backwards to the lines of activity set off earlier.
- Finally, the third level is termed microstructure, realized as moment-by-moment texture across visual and audio modes, e.g., how devices such as cutting patterns in action sequences and ongoing dialogues "cooperate" to function in a larger meaning unit.

Although Bordwell's analysis of the hook lays the emphasis mostly on the devices operating at the third level of microstructure, he nevertheless states that one can study film coherence at any of these levels; the devices of scene transition can also play roles in segmenting a structure at the two higher levels. For instance, the end of a scene sometimes simultaneously symbolizes the end of a plot or an act. The analysis of *2046* put forward in this paper will highlight precisely such cross-level interactions. It will show just how audiovisual cohesion at the third level can be used to tighten the unity of a film with loose plot coherence in its macrostructure.

The remainder of this section will review Bordwell's discussion of the concrete devices of audio and visual hooks and their general functions and constructions at this micro-level of organization. This is to pave the way for the next step, namely, for particularly unraveling how Bordwell's descriptive scheme can actually be strengthened by constructing cohesive chains.

According to Bordwell, the hook is said to have been first extensively used by the German director Fritz Lang for two purposes: accelerating the pace of a film, and strengthening the core motifs of a scene (cf. Burch 1991: 3–31). In this perspective, the overall function of the hook is to "create a more overt narration . . . This sort of strategy can work in classical films because the beginnings and the endings of scenes are typically the most overt portions and this overt narration can build up a sense of cleverness or resourcefulness or sophistication" (Bordwell 2008). Furthermore, the hook can also be used to enhance the suspense of a film – when important story events need to be concealed for such an effect, the hook is often used to create ellipses of these events.

With regard to the construction of texture in scene transitions, Bordwell maintains that scenes can generally be hooked together through two main types of cohesive construction: (1) when a character's movement/action at the end of one scene is connected directly to the actions at the beginning of the next scene, exactly as in the transition in *The Birds* sequence where Melanie is seen walking continuously across two scenes, and (2) when the first kind of transition is not the case, namely, no match-on-action is employed across the transition. If the second type of construction is selected, certain devices, i.e., the hook, must be mobilized to function as semantic connections across the cut.

Take Bordwell (2008)'s analysis of the film *National Treasure* for instance. Bordwell first demonstrates how a sound element or verbal text can hook to an image, or vice versa, how an image can hook back to a sound and to a verbal text, etc. His analysis shows that the most common type of scene transition realized in this film is a cross-modal one: realized as "the hook from a sound/dialogue to an image." An example is displayed in Figure 9. The context before these images is that the man and the woman are trying to bring out the cipher on the back of the declaration that they found; and thus they squeeze on lemon juice and blow fervently on the paper. In the first two images shown in the figure, the man said: "we need more juice" and then the woman says: "we need more heat," this scene is cut to a bowl of lemons, then to a drawer where a hair dryer lies. After the cut, the man and the woman continue to coax the cipher out of the paper. In this example, the hooks "juice" and "heat" operate to connect dialogues and visual images across these different scenes.

Bordwell's analysis also demonstrates that the hook need not be literal, linking a continuing narrative: very often, it can be metaphorical or misleading in order to create a playful narration. An example of this used in his essay is displayed in Figure 10, extracted from the same film. The man (Riley) in the first image is seen checking on some video installations and says "Game on"; this is cut to the second image, a digital readout counting down from 3 to End. The third image then shows that the digital number is not a timer in the man's equipment; rather, it belongs to a microwave oven in which the second man (Ben) is preparing food. In this example, what the viewers take as a continuity editing within a scene is actually a hook between two different scenes. This kind of disorientation brought about by a misleading hook also functions to invite the viewers to pay more attention to just how the semantics in two scenes fit together.

More importantly, Bordwell's functional description of cinematic cohesion highlights just how viewers' dynamic process of inference is realized – the meaning construction of this "playful, misleading" hook relies precisely on viewers'

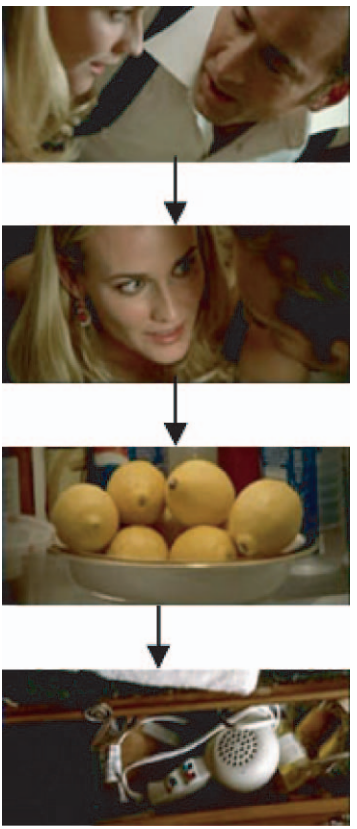


Fig. 9: Dialogue transition – sound/dialogue hooks to an image

ongoing revision of narrative interpretation as the film unfolds. Functional analysis in this fashion is conducted in a problem-solving manner, one typical application of the neo-formalist approach to film stylistics. In this particular case, when exploring a scene transition Bordwell first describes the general cohesive function of the hook for creating texture across scenes, illustrates the formal patterns of the hook (e.g., an image hooks to a sound, a dialogue hooks to an image, etc.) by examining a particular film, and then further discusses its historical and conventional uses.

To this point I have been reviewing the neo-formalist approach to cinematic texture and scene transition. The rest of this section will show just how the problem-solving description can actually be more systematically and effectively represented through patterns of cohesive chains.

Fig. 10: Example of "misleading hook"

Based on the same principle of chain construction illustrated in Section 2.2. Figure 11 maps out the chain patterns established from Bordwell's example extract in Figure 10. In this extract, five chains are isolated: "the first setting at National Archive," "Riley," "equipment," the other man "Ben," and the second setting at Ben's studio. The formation of the chain pattern explicitly shows a dominant chain functioning as a "bridge" connecting the two scenes. This bridging "equipment" chain is established by a few *similar* objects ("digital numbers," "video installation," and "microwave oven") falling broadly into the same object category of "equipment." In other words, the relation between these elements is not strictly "co-reference" but another type of cohesive device, comprising a semantic relation that realizes ties of similarity. Drawing on this linguistic notion, this kind of semantic relation is established through two types of *similarity* ties: a

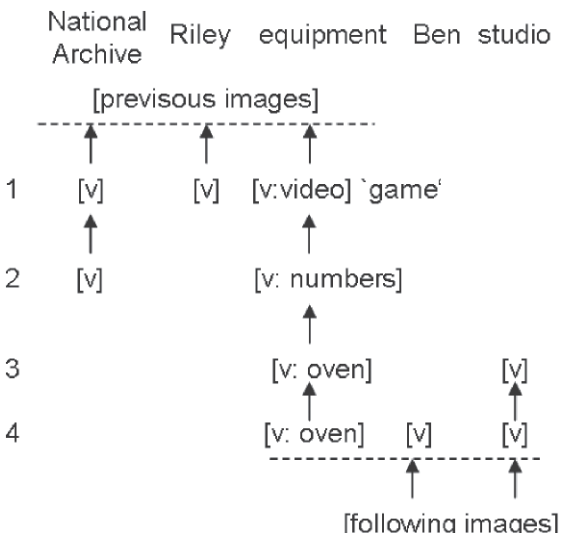


Fig. 11: Primary chain patterns built from the extract in Figure 10

part-whole relation between two elements termed *meronymy*, and *hyponymy*, which refers to elements under the same broader classification.[1]

In the present case, the relation between "digital numbers" and "video installation" is meronymy, denoting a part–whole relation, while "video installation" and "microwave oven" are co-hyponyms, considered as falling under the same classification, electronic equipment.

Furthermore, comparing this example to the cohesive chains in The Birds, the chain formations of the two are similar in that the viewers are taken across two different scenes through the continuity of dominant object/characters – Figure 7 shows that the dominant character/object chains "Melanie" and "birds" take the viewers across the two settings and in Figure 10 we also see a longer chain "equipment" bridging through the two different scenes.

To be more specific, although the two different ways of connecting scenes maintained by Bordwell are (1) *with* continuous movements of characters (e.g., the extract in *The Birds*) and (2) *without* (e.g., the extract in *National Treasure*), the comparison of the chain patterns between the two extracts nevertheless shows that the paths for the viewer's interpretation in this respect are not totally

1 For the full treatment of the application of the semantic relations of hyponymy and meronymy to moving images, see Tseng (2008) and Chapter 4 in Tseng (2009).

different. The discourse strategies used for signaling the coherent flow of narrative to the viewers can actually be analogous and it is the *continuity* of characters' and objects' identities that plays the significant role in these cases and in these paths.

Hence, the comparison precisely reflects what is otherwise not easily brought to light through a problem-solving description – in this case, the systematic construction of cohesive chains uncovers the similarity of viewers' interpretive paths in scene transitions although the two extracts actually deal with substantially different kinds of story events.

Finally, as Bordwell specifically points out in his essay, there are several possibilities for cross-modal hooks such as a sound hooking to an image or an image hooking to a sound. In this respect, what is particularly useful of cohesive chains is the potential for showing such cross-modal-ness, namely, how a scene hooking mechanism, i.e., a bridging chain, is actually cross-modally realized.

4 Audiovisual texture and narrative interpretation in *2046* (2004)

4.1 Brief description of scenes

In this section, the analytic methods introduced above are applied to describe how one of the main narrative effects of the film *2046* is achieved. As already mentioned, this film unfolds along a few non-linear temporal settings, ranging from scenes in the fictional world created by the protagonist Chow, scenes about Chow's relationships with two women in Hong Kong from 1966 to 1970, scenes of Chow's memories about his relationship in Singapore in 1963 and 1969, to some pieces of brief scenes depicting Chow's older memories.

Figure 12 maps out how the temporal transitions of these scenes operate as the film unfolds in time. Through illustrating the diagram, I will provide a (necessarily brief) summary of these plotlines.

- *Scene A*: The film begins with a scene in the far future. As the story begins, a Japanese time traveler Tak (Kimura Takuya) narrates his long train ride returning from 2046 and a woman (Wong Faye) he fell in love with. This scene ends with an intertitle (*a* in the diagram) "All memories are traces of tears."
- *Scene B*: This is the scene where Chow (Tony Leung) is first presented. It shows Chow's intention to take a woman, Su Li-Zhen (Gong-li) with him back to Hong Kong but Su rejects his request. This scene ends with Chow's

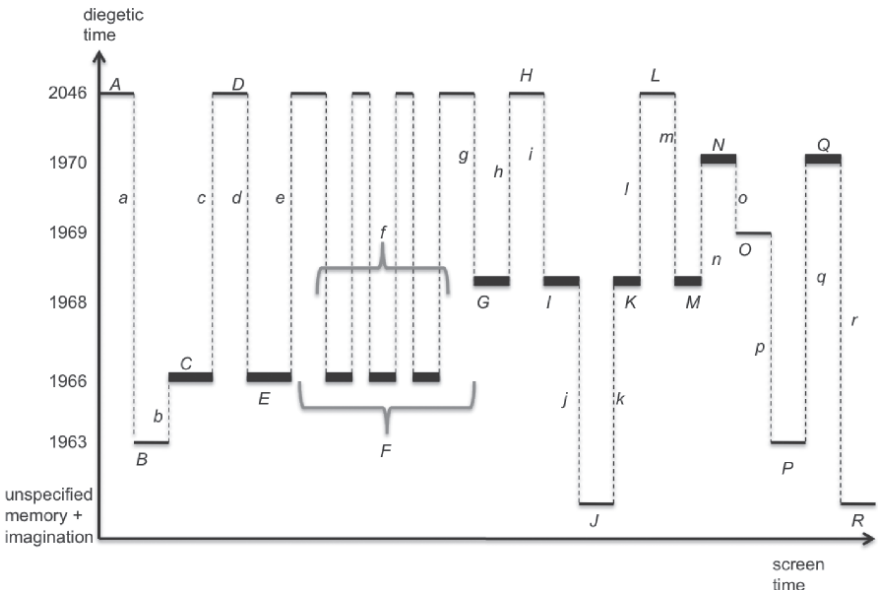


Fig. 12: Transitions of the major temporal settings in 2046. Each main sequence of different temporal settings is specified as A–Q, and the scene transitions are shown as the dotted lines a–q. The scenes of F and their transitions f comprise a bunch of rapidly alternating scenes cut across the fictive settings of 2046 and Chow's writing scenes in 1966

off-screen narration about his return to Hong Kong from Singapore. Hence, the transition *b* from scene B to scene C in Hong Kong is realized by Chow's continuing narration. But the temporal setting in scene B remains unspecified. It is retroactively clear only in the later scenes (scene O and scene P) that it takes place around 1963.

– *Scene C*: The scene begins with Chow's off-screen narration about how he moves into room 2046 in a hotel and the people he encounters. His narration substantially anchors the visualtrack and realizes high degrees of cohesive harmony across the visualtrack and verbal text (Tseng 2008). In Figure 12, scenes C, E, some of F, G, I, K, M, N, and Q are specified with bold lines for demonstrating how these scenes are arranged into a major chronological plotline, while other scenes function more like devious narrative routes that depict Chow's outlandish imaginations and some non-chronologically arranged scenes of his memories.
– *Scene D*: The transition between scenes C and D, which is again the same fictional setting as scene A, is realized by Chow's narration about how he starts writing a story called *2046* and envisions the women and men

around him as characters in this novel. This cues the viewers to relating retroactively the diegesis of scene A as the story Chow has been writing. In effect, throughout this film, the device of Chow's verbal depiction is substantially used as the bridge in scene transitions (e.g., in *b, c, d, e, f, h, i, j, k, l, m, o, p*) to orient the viewers across the settings of the Hong Kong hotel, 2046 and his memories.

- *Scenes E and F*: After scene C, scene E functions as the further development of the same, chronological storyline. It centers on Chow's relationships with two women: Bai Ling (Zhang Ziyi) and the daughter of the hotel owner Jinwen (Wong Faye). He develops romantic affection for the latter and novelizes himself and Jinwen as Tak and an Android (also played by Wong Faye, as already shown in scene A) in his story. Scenes in F are composed of a few rapid intersections of alternating scenes of Chow's imagined world and some scenes showing Chow writing in front of his desk in the hotel room.
- *Scene G*: Scenes in F end with a longer scene in the 2046 setting and are followed by an intertitle (*g* in the diagram) "24 December, 1968." Scene G then shows Chow and Jinwen having dinner Christmas Eve, 1968.
- *Scenes H, I, J, K, L, M*: These scenes are composed of complex layers of settings across Hong Kong around 1968, Chow's imagination, and his memories. These chronologically unrelated scenes in the visualtrack are tied together by Chow's continuous verbal depiction across the scene transitions.
- *Scenes N, O, P*: Scene N depicts a new temporal setting in 1970, which is specified by an intertitle "after eighteen months" (*n* in the diagram). In scene N, Chow meets with Bai Ling and this encounter reminds him of their past relationship in Singapore. This then triggers the next two scenes, O and P, depicting Chow's search for Su Li-Zhen in 1969 and his previous relationship with her in 1963 in Singapore.
- *Scene Q*: The scene transition from P to Q is realized by an intertitle "When peony blooms, she stands tall and goes away. Does she mean 'no' or 'yes'?" Unlike the previous intertitles, it does not specify the temporal relationship between the scenes and this is followed by scene Q showing Chow's interaction with Bai Ling in Hong Kong, presumably not too far from their last encounter depicted in Scene N. This scene ends with Chow's leaving Bai Ling. The transition *r* from scene Q to scene R is again an intertitle "He didn't return back. It was as if he'd boarded a very long train. Heading a drowsy future through an unfathomable night."[2]

2 The English translation of the intertitles are provided by the English version of the movie.

– *Scene R*: The last scene of this film comprises some black-and-white images of Chow in a taxi. Their temporal settings are not specified, but black-and-white images suggest that the scene could be in Chow's memory.

The wave-like diagram in Figure 12 visualizes a clue, revealing why this film has been regarded as puzzling (Bettison 2010) since its release – the complex intersection of narrative lines constructs an interpretation path that is quite demanding for the viewers to follow. As the above description illustrates, the scene transitions (*a*–*q*) in this film employ two major types of film devices: (1) intertitles, e.g., transitions *a, g, n, q* and (2) continuity of Chow's off-screen narration, e.g., transitions *b, c, d, e, f, h, i, j, k, l, m, o, p*.

The following cohesion analysis will unravel just how these devices realize some discourse strategies, which function to orient as well as to disorient the viewers along these intertwining story threads.

4.2 Cohesive chains of the scene transitions

As described above, the film begins in the year of the film's title, and the sequence constructs a futuristic atmosphere with an artery of railtracks, tunnels, and fast moving trains. The Japanese man Tak's voiceover accompanies images in the train and female androids. Figure 13 transcribes the transition from scene A to scene B, taking place from image 4 to image 5 in the figure. This move lacks any expository markers for orienting the viewers in the interpretation path. The verbal text in the intertitle does not help the viewers to determine the spatio-temporal context of the next scene. The only verbal clue is "memory" in the text and this could signal the next scene as a part of someone's memory – however, no cue in scene B, (in which Chow and Su are first presented in the film,) indicates this connection.

Figure 14 maps out the cohesive chains established from the extract in Figure 13. The pattern not only shows a lack of bridges across the two settings, thus leaving spectators disoriented in the narrative inference path, but also a lack of specific cues for identifying the second setting – it remains as a generic place until the end of this scene, where Chow's voiceover finally mentions Singapore.

The next transition from scene B to scene C, transcribed in Figure 15, is realized, unlike the first one, through the explicit verbal cues in Chow's voiceover in image 3, following an insertion of a shot showing the fast-moving trains in 2046. The cohesive chains of this extract are shown in Figure 16. The explicit identification can be seen from the cross-modal elements of the setting chains, namely, [v] "Singapore" in the first setting's chains and the [v] "Hong Kong," [v] "Kowloong"

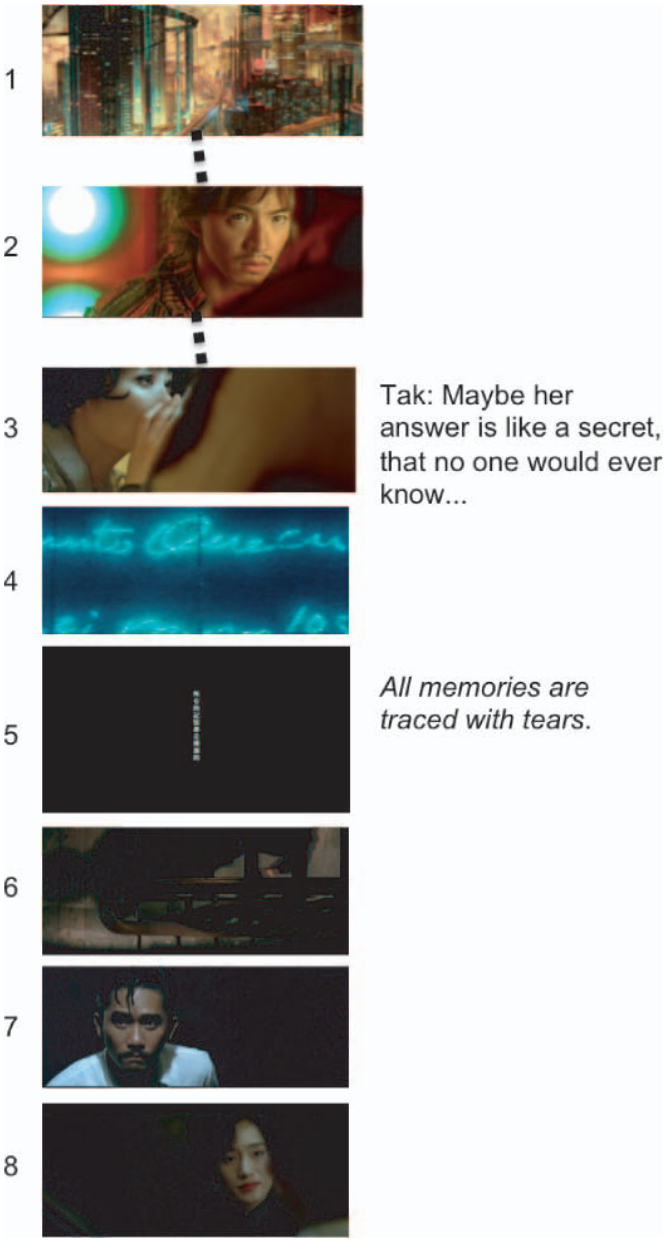


Fig. 13: Transition from scene A to scene B

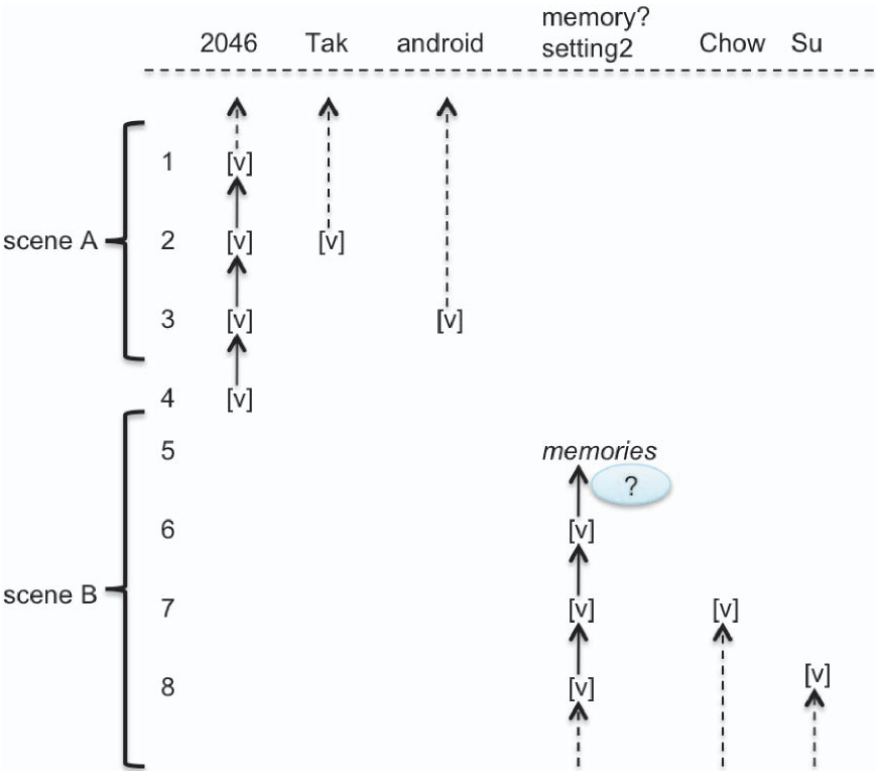


Fig. 14: Cohesive chains of the transition a: from scene A to scene B. Dotted arrows refer to the continuity of chain constructions from previous sequences or scenes

in the next setting's chain. Although an image of the futuristic train (image 2 in the figure) briefly carries the viewer to the setting of scene A, the following images and Chow's description of his Hong Kong life soon guide the viewers to the third setting of the film. The chains established in this extract demonstrate how, despite the short detour of the 2046 setting, Chow's chain is constructed across the transitions. Hence, unlike chain patterns in Figure 14, here we see a robust bridge, with the Chow chain carrying the viewers across scenes.

Scene C ends with the setting where Chow is writing in front of his desk and explains in a voiceover how he is starting to write a novel called *2046*. The transition from scene C to scene D is displayed in Figure 17. Similar to the previous transition, Chow's voiceover also explicitly leads the viewers to the next setting. Image 2 and image 3 depict the scene transition in this extract and here we can see how the identity tracking is realized cross-modally again in this context. Not

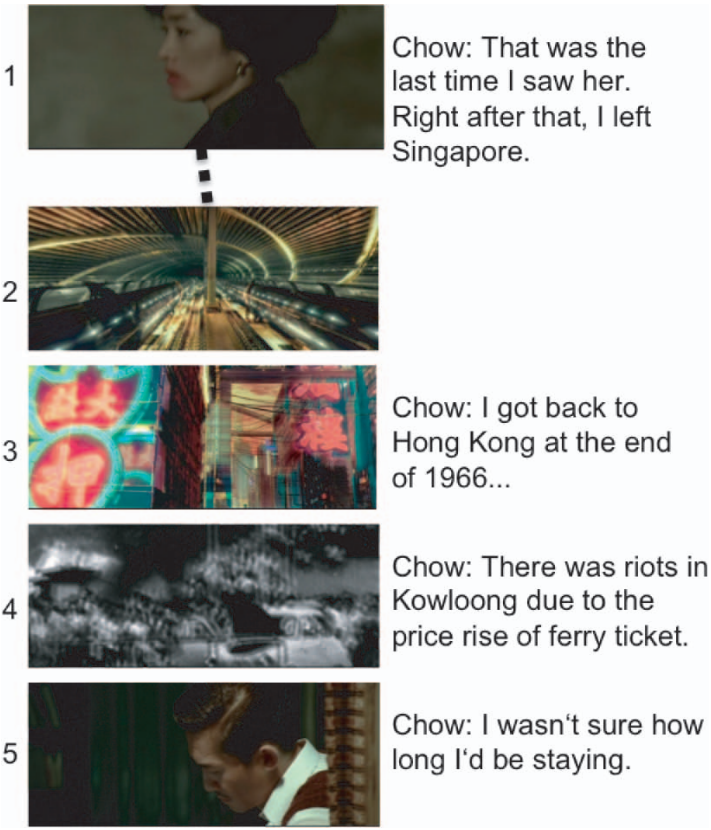


Fig. 15: Transition from scene B to scene C

only does Chow refer to himself several times in the voiceover "I stopped . . . ; I was . . ." when he is seen in the visualtrack in image 1, but also in image 2, his hand and his action of writing is accompanied by the verbal text "I was writing. . . ."

The cohesive chains of this extract are shown in Figure 18. Like the pattern in Figure 16, the Chow chain is established throughout the two settings and the cohesive element in the 2046 chain, i.e., [v] "2046," also demonstrates cross-modal links for explicitly identifying the setting that the viewers should be oriented to. This kind of pattern formation, namely, with the Chow cohesive chain clearly bridging across scenes and with cross-modal cues for scene identification, is actually employed again and again in the following transitions *d* (scene D to scene E), *e* (scene E to scene F) and *f* (a few alternating scenes of 2046 and Hong

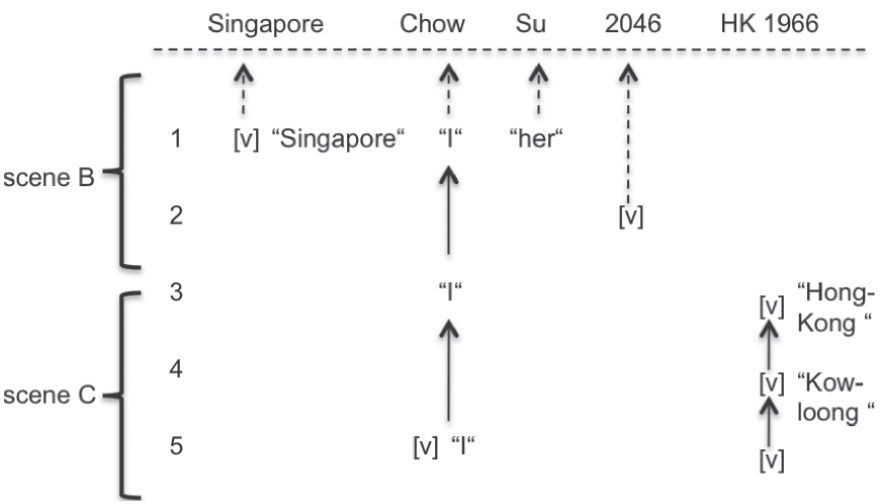


Fig. 16: Cohesive chains of the transition b: from scene B to scene C. Dotted arrows refer to the continuity of chain constructions from previous sequences or scenes

Fig. 17: Transition from scene C to scene D

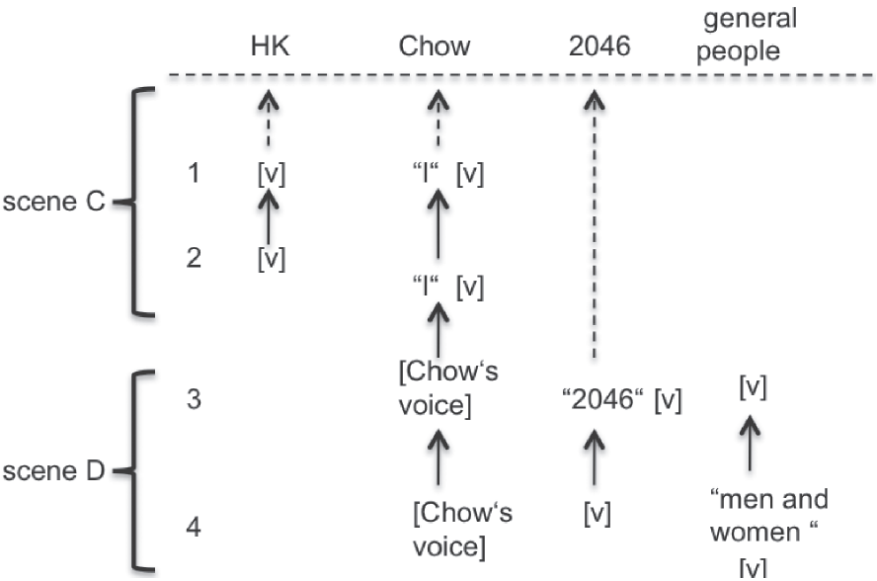


Fig. 18: Cohesive chains of the transition c: from scene C to scene D. Dotted arrows refer to the continuity of chain constructions from previous sequences or scenes

Kong). As for the transition from scene F to scene G, it is verbally identified by the intertitle "December 24, 1968." This cohesive pattern is similar to the transition *n*, which will be demonstrated in the next formation of cohesive chains.

In *2046*, what seems particularly puzzling, apart from scene transition *a*, is the transitions from scene I to scene M, transcribed in Figure 19. Scene I ends with Chow walking on the street after intending but failing to revise the ending of his novel from a sad ending to happy ending. After he calls the taxi on the street (image 1), scene I jumps abruptly to scene J which is composed of black and white images (image 2 and 3) set in a taxi. Scene J presumably visualizes Chow's memory about his past relationship and this narrative inference is supported by his voiceover ". . . Some years ago I had a happy ending in my grasp but I let it slip away . . ." Chow's voiceover continues to lead us to the next setting where he bumps into Lulu again, a woman he once knew and also novelized in his *2046* fiction. A clear point-of-view construction from image 4 to image 9 (images 4, 5, 6: scene K, images 7, 8: scene L, image 9: scene M) shows how Chow first watches Lulu in a bar (images 4, 5, 6) and then envisions her as an android again in his fictive world (images 7, 8, 9). Finally, starting with an intertitle explicitly referring to the temporal setting as "eighteen months later," scene N depicts Bai Ling calling Mr. Chow and her meeting with Chow for a drink.

Fig. 19: Transitions from scene I to scene N

The cohesive chains constructed from this extract are displayed in Figure 20. The chain pattern across these fast transiting scenes surprisingly shows no dramatic difference from those patterns in *The Birds* (Figure 7) in that at least a long continuing protagonist chain is constructed throughout the extracts and functions to bridge different scenes. Furthermore, although Chow's voiceover does not specify the setting of the visual images and the links in the setting chains are not

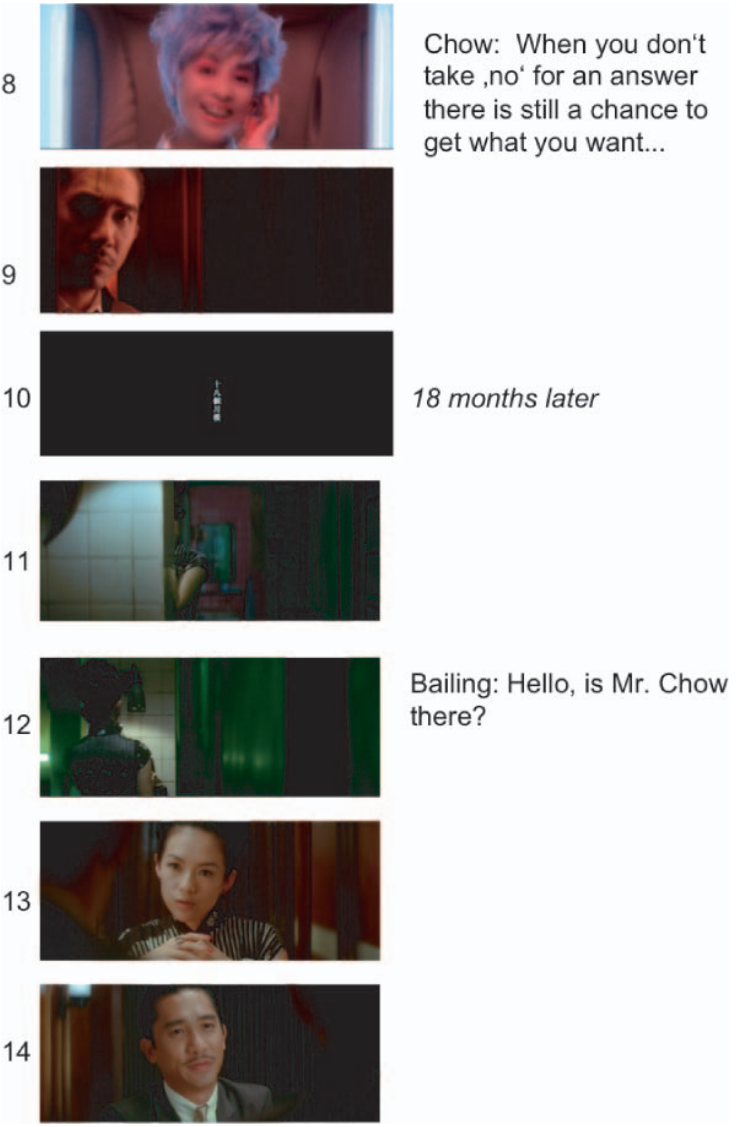


Fig. 19: (Cont.)

cross-modal as are those in the two previous patterns, the identities of the settings and characters in this extract are actually not newly presented. Rather, the dotted arrows in the chain pattern refer to the fact that these characters and settings have been introduced in previous scenes and refer to the repetition of visual

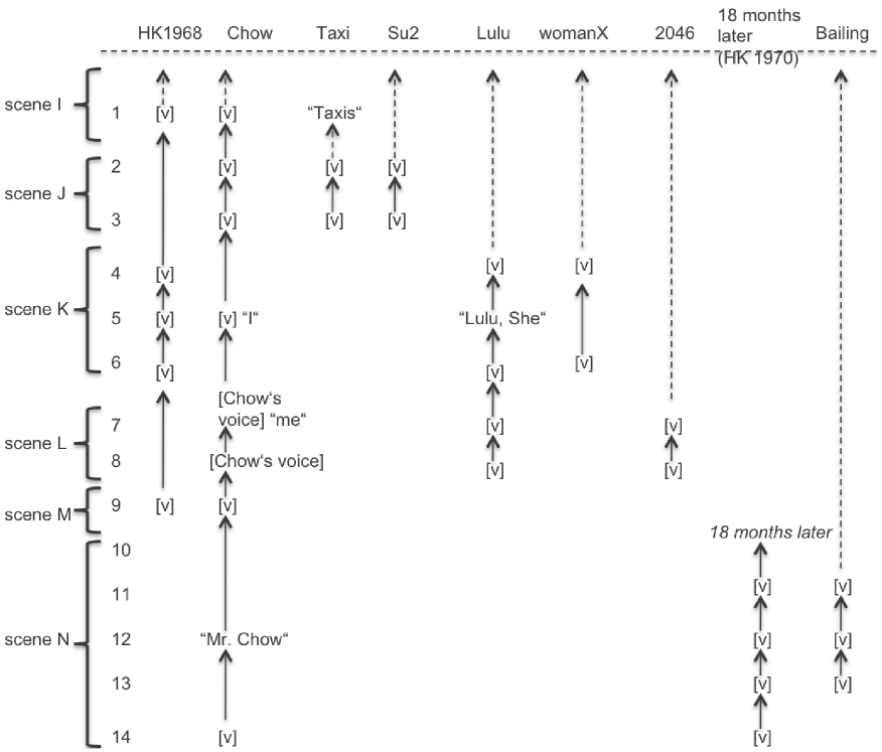


Fig. 20: Cohesive chains of the transition i–n: from scene I to scene N. Dotted arrows refer to the continuity of chain constructions from previous sequences or scenes.

images, for instance, Lulu as an android in a 2046 train, Su Lizhen (played by Maggie Cheung) with Chow in the taxi in the black and white images, have been presented and identified earlier in the film. The viewers should be able to recognize them again since their reappearances are frontal and explicit in these images.

The cohesive chains then untangle a complex route of narrative construction. What is often puzzling is NOT the question of *who* and *where*, namely, which characters or what settings in the film we are viewing; rather, *why* they are here. In other words, the uses of visual logical cues, such as the black and white images which could possibly indicate a character's reflection into the past, or the point of view shots constructed in images 5–9 suggesting Chow's imagination, are not dominant and salient enough as verbal texts to guide the viewers along the rapid transitions in such a short time.

In contrast to these fast transiting scenes from scene I to scene M, the transition from scene M to scene N is clearly guided by an explicit intertitle "eighteen

months later," which finally leads the viewers back to a narrative line with a clearer temporal setting. As we can see in the chain pattern, after image 10, another setting chain, i.e., Hong Kong 1970, is established following the ending of the first setting chain, of Hong Kong 1968.

5 Discussion and conclusion

The above cohesion analysis of *2046* uncovers a significant aspect of the narrative complexity of this puzzling film and I think that it, further, raises some empirical questions.

As the chain patterns in the previous section indicate, except for the first transition, the audiovisual texture in this film is actually well constructed. The discourse strategies of cohesion at work set the viewers onto the preferred narrative comprehension path; in particular, verbal texts or visual cues (sometimes both) are often employed to guide the viewers to identify characters and settings across different scenes. The comparison of chain patterns between *2046* and other films such as *The Birds* argues against the claim that *2046* demands totally different processes of comprehension. Although logical cues between scenes are not always available or salient (e.g., the scenes analyzed in Figure 19/Figure 20), the audiovisual texture nevertheless operates to compensate for the unpredictable temporality. In brief, setting out the cohesive identification cues mobilized in these sequences clearly shows how the film, despite its unconventional sequencing, is actually largely cohesive.

This analytical result further brings to light an interesting cognitive phenomenon illustrated by Bordwell:

> Reading, notes Barthes in *S/Z*, involves forgetting. So does viewing ... As practical psychologists, our filmmakers know that we'll construct a diegetic world chiefly through landmarks, not fine details of setting. They know that we'll move rapidly from items of appearance and behavior to inferences about character beliefs and traits. And they know that under the clock, we're likely to overlook stylistic features. (Bordwell 2007: 143)

In a certain respect, this explains why viewing *2046* is widely regarded as demanding and puzzling: despite the explicit use of visual stylistic cues to signal the relation between these intersecting storylines, e.g., the black and white images symbolizing the settings in Chow's memory, the repetitive patterns of story actions (Bettison 2010: 170–172) strengthening the parallels between Chow's life and the *2046* story, it's nonetheless uncontentious that for first-time viewers, constructing narratives and making inferences is no easy task.

Another related issue is the functional gap between the verbal cues and visual cues for orienting viewers across different scenes. This gap is particularly prominant in *2046* in that the verbal texts realized by Chow's voiceover and intertitles are often safer guides for the viewers than the visual patterns of repetitive settings and actions. This raises the question as to how different modes, although performing the same function as cohesive identification cues, achieve different degrees of coherence in viewers' narrative construction. It is in this light, I believe, that the difference of the dominance between verbal and visual cues for narrative construction requires further empirical investigations.

In sum, in this paper I demonstrated a discourse method for analyzing filmic cohesion, that is, audiovisual texture mobilized during the viewers' comprehension process. I also applied this method at the empirical level, showing how it can unravel narrative construction across scenes. The focus of analysis was anchored in how dominant narrative elements such as characters, objects and settings are coherently presented and represented as a film unfolds. Moreover, I also elucidated how the analytical method can be seen as complementing and strengthening the problem-solving neo-formalist approach by providing a more systematic analysis. Analysts who intend to investigate meaning-making in film need methods for articulating accounts that are able to elucidate just how the construction of filmic devices directs, guides and constrains the viewer's contextualized interpretation (cf. Bordwell 1989); and I hope to have shown here that filmic cohesion can precisely contribute to these crucial components.

References

Bateman, John A. 2007. Towards a *grande paradigmatique* of film: Christian Metz reloaded. *Semiotica* 167(1/4). 13–64.

Bettison, Gerry. 2010. Happy together? Generic hybridity in *2046* and *In the mood for love*. In Warren Buckland (ed.), *Puzzle films: Complex storytelling in contemporary cinema*, 167–186. Chichester: Wiley-Blackwell.

Bordwell, David. 1985. *Narration in the fiction film*. Madison, WI: University of Wisconsin Press.

Bordwell, David. 1989. *Making meaning. Inference and rhetoric in the interpretation of cinema*. Cambridge, MA: Harvard University Press.

Bordwell, David. 1997. *On the history of film style*. Cambridge, MA: Harvard University Press.

Bordwell, David. 2006. *The way Hollywood tells it: Story and style in modern movies*. Berkeley & Los Angeles: University of California Press.

Bordwell, David. 2007. *Poetics of cinema*. London & New York: Routledge.

Bordwell, David. 2008. The hook: Scene transitions in classical cinema. http://www.davidbordwell.net/essays/hook.php (accessed 12 August 2012).

Burch, Noel. 1991. *In and out of synch: The awakening of a cine-dreamer*. Aldershot: Scolar.

Eder, Jens. 2008. *Die Figur im Film: Grundlagen der Figurenanalyse*. Marburg: Schüren.

Eder, Jens. 2010. Understanding characters. *Projections* 4. 16–40.
Eggins, Suzanne. 1994. *An introduction to systemic functional linguistics*. London: Pinter.
Eisenstein, Sergei. 1969. *Film form: Essays in film theory*. PLACE OF PUBLICATION: Harvest.
Halliday, Michael A. K. & Ruqaiya Hasan. 1976. *Cohesion in English*. London: Longman.
Halliday, Michael A. K. & Christian M. I. M. Matthiessen. 2004. *An introduction to functional grammar*, 3rd edn. London: Edward Arnold.
Hasan, Ruqaiya. 1984. Coherence and cohesive harmony. In James Flood (ed.), *Understanding reading comprehension: Cognition, language, and the structure of prose*, 181–219, Newark: International Reading Association.
Martin, James R. 1992. *English text: Systems and structure*. Amsterdam: Benjamins.
Metz, Christian. 1974. *Film language: A semiotics of the cinema*, Michael Taylor (trans.). Oxford & Chicago: Oxford University Press & Chicago University Press.
Palmer, R. Barton. 1989. Bakhtinian translinguistics and film criticism: The dialogical image? In R. Barton Palmer (ed.), *The cinematic text: Methods and approaches*, 303–341. New York: AMS Press.
Schegloff, Emanuel A. & Harvey Sacks. 1973. Opening up closings. *Semiotica* 8(4). 289–327.
Smith, Murray. 1995. *Engaging characters: Fiction, emotion, and the cinema*. Oxford: Oxford University Press.
Thompson, Kristin. 1988. *Breaking the glass armor: Neoformalist film analysis*. Princeton, NJ: Princeton University Press.
Tseng, Chiaoi. 2008. Cohesive harmony in filmic text. In Len Unsworth (ed.), *Multimodal semiotics: Functional analysis in contexts of education*, 87–104. London: Continuum.
Tseng, Chiaoi. 2009. *Cohesion in film, and the construction of filmic thematic configuations: A functional perspective*. Bremen: University of Bremen dissertation.
Tseng, Chiaoi & John A. Bateman. 2010. Chain and choice in filmic narrative: An analysis of multimodal narrative construction in *The Fountain*. In Christian R. Hoffmann (ed.), *Narrative revisited: Telling a story in the age of new media*, 213–244. Amsterdam: John Benjamins.

Bionote

Chiaoi Tseng (b. 1976) is a staff scientist at Bremen University ⟨chiaoi@uni-bremen.de⟩. Her research interests include film analysis, multimodal discourse and genre, and techniques of automatic processing for high-level narrative analysis. Her publications include "Multimodal narrative construction in Christopher Nolan's *Memento*: A description of method" (with J. A. Bateman, 2012); "Revisiting cinematic authorship: A multimodal perspective" (with J. A. Bateman) and *Cohesion in film: Tracking film* (in press).