Coherence and Cohesive Harmony in Filmic Text

Chiaoi Tseng

Abstract

This chapter is concerned with the ways in which filmic cohesion contributes to the creation of text coherence and seeks to expand the linguistic approach developed by Hasan (1984; 1989) so as to apply to multimodal discourse. First, a synopsis of the construct of cross-modal cohesive chains is provided; this is based on the framework of multimodal cohesive reference mapped out in detail by Tseng (2008). Second, the interaction of filmic cohesive chains is modelled to see if cohesive harmony of a filmic text can be examined along the same lines as that in language text. The result of this study shows that examining cohesive quality of a multimodal text needs to take into account chain interaction in two aspects: (1) mono-modal chain interaction based on Hasan's method and (2) cross-modal chain interaction drawing on synergistic ties across different modes. Finally, some thoughts are given to the possible contributions that cohesive chains and cohesive harmony can make to the further development of filmic discourse analysis.

1 Introduction

Within film there is a complex interaction of co-occurring modalities, e.g. words, images, sounds, colours, actions, etc. that combine and cohere to create meanings. This chapter will propose one method for investigating just how this meaning creation takes place. Methodologically, I explore to what extent some analytic constructs from systemic functional linguistics can be extended to apply within and across filmic modes; in particular, I will investigate the contribution made by Hasan's (1984; 1989) notion of coherence and cohesive harmony to the construction of meaning in film.

Briefly speaking, this chapter aims at proposing a tool for filmic analysis, which is essential and has not been fully developed yet, particularly from the following perspectives— First, multimodal frameworks based on social semiotic theory have been substantially developed in recent years (cf. Kress & van Leeuwen, 2001; O'Halloran, 2004a; Ventola et al., 2004; van Leeuwen, 2005; Baldry & Thibault, 2006; Bateman, *forthcoming*-a); however, the pursuit of common investigative tools for multimodal analysis often poses difficulties. As

Macken-Horarick (2004, p. 6) points out, "The pursuit of a common framework for analysing multimodal texts produces two kinds of awareness… (a) awareness of a 'lack of fit' between categories of one mode applied to another; and (b) awareness of the deconstructive power of this kind of analysis, which reveals gaps, silences and, surprisingly, providential riches in transmodal analysis". For strengthening the analysis of interaction between modes, we still need more devices for *systematically* describing the *cross-modal realization* of meaning. Hence, this chapter aims to propose one of such 'bridging devices'. I propose that examining how different semiotic modes work together to establish a *coherent* semantic edifice is an effective way to elucidate the cross-modal interaction and its contribution to the total meaning construction; and for this, Hasan's linguistic methods for examining text coherence are proposed as appropriate tools for application since her aim is precisely to show how units in a language text are brought together to cohere.

Second, film theorists to date have attempted to provide useful tools for film interpretation and analysis but have not been able to formulate a well-rounded method that allows the investigations of meanings in filmic texts across different types and genres. This difficulty stems from their misrecognition that a corpus of filmic texts for an analytical purpose should be drawn based on the industrial/contextual process (cf. Mittell, 2001; 2004). Hence most film theorists tend to develop analytical approaches specifically for fiction films (cf. Bordwell, 1984; 1989), for documentary films (cf. Nichols, 2001), for television programs (cf. Mittell, 2004), and so on. It has been taken for granted that filmic texts under those different 'contextual headings' must be vitally different and there should be discrete systems for each merely because they are products of distinct mediums with different industrial imperatives and audience practices. Thus, when it comes to the intertextual comparison across filmic texttypes, i.e. the compatibility of a type of filmic texts in different mediums, or the issue of filmic genre, most analytical approaches become insufficient: and this is one reason why the notion of film genre has been a perennial conundrum that has plagued the realm of film for decades (cf. Tudor, 1973; Gledhill, 1985; Neale, 2000). On this account, the method for examining filmic coherence proposed in this chapter, alongside the previous development of methods for filmic analysis carried out by some researchers from the social semiotic perspective (cf. van Leeuwen, 1985, 1991, 1996; O'Halloran, 2004b; Baldry & Thibault, 2006; Bateman, *forthcoming*-b) can be seen as another significant tool for investigating meanings in filmic texts of various genres, and these analytical tools are imperative steps towards the future development of a more cogent model for film genre analysis and comparison.

In the following sections, I will first review Hasan's conception of coherence and cohesive harmony in language text, and then examine if filmic cohesive chains and chain interaction based on Hasan's principle can be used to investigate the cohesive quality of filmic texts. Through an analysis of a documentary film, I will propose a complementary approach to cross-modal cohesive harmony which lays bare the synergistic ties between units across different modes. Finally, I will give some thought to how the construct of filmic cohesive chains and cohesive harmony can contribute to the further development of more theoretical tools for filmic discourse analysis.

2 The notion of cohesive chains and cohesive harmony

Hasan, in her 1984 and 1989 work, proposes several methods for investigating how cohesion contributes to the creation of coherence in a language text. Her approach consists of generating lexical chains of a text on the basis of lexical cohesive ties (Halliday & Hasan, 1976), followed by an investigation of the lexical tokens that realize two types of lexical chains: (1) *identity chains* consisting of identities of people, places and things and (2) *similarity chains* whose components are concerned with actions or qualities of the participants (Hasan, 1989, p. 84). Hasan describes these cohesive chains as "threads of continuity running through the text" and asserts that each cohesive chain "supports and refines the domain of meaning for the others" (p. 89). The construct of both types of cohesive chains then results in the distinction of lexical tokens into *relevant tokens*, which enter into identity chains or similarity chains, and *peripheral tokens*, which do not enter into any kind of chain. Consider the following text:

1. Once upon a time there lived a young frog.

2. His name was Sniffy.

3. He often sat upon a very soft lily pad,

4. probably because he liked flowers.

5. One day, Sniffy was scrubbing his toes in the pond.

6. He felt something moving in the water.

7. It was his friend, Fluffy.

Nine cohesive chains are found in this text: the chains a—e are identity chains and the chains f—i are similarity chains, displayed as follows.

Identity chains:

a	b	c	d	e
1. once upon a time 5. one day	1. frog 2. his name, Sniffy	3. lily pad 4. flowers	5. pond 6. water	6. something 7. it, friend, Fluffy

	3. He 4. he 5. Sniffy, his 6. he 7. his			

Similarity chains:

f	g	h	i
4. like 6. felt	3. sat 6. moving	2. was 7. was	1. young 3. soft

When we examine the lexical tokens in this text, the tokens incorporated in the nine chains are relevant tokens and the remaining lexical tokens in the text, e.g. "scrubbing", are peripheral tokens according to Hasan's definition.

However, cohesive chains as such merely consist of a succession of message constituents rather than whole messages; but texts consist of a concatenation of messages. Thus, Hasan points out that if cohesion is to demonstrate what makes a text coherent, there has to be some way of examining the contribution of message components to messages as wholes. To quote Hasan (1989, p. 91):

> Although the chains go a long way towards building the foundation for coherence, they are not sufficient; we need to include some relations that are characteristic of those between the components of a message. This is the relation that I refer to as **chain interaction**… A minimum requirement for chain interaction is that *at least two members* of one chain should stand in the same relation to *two members* of another chain.

Building on Hasan's principle, the chain interaction of the *frog* text is displayed in Figure 1. There it can be seen that only three out of the nine chains are involved in the interaction: chain *a* is connected to chain *b* by the relation of circumstance and chain *b* to chain *f* by senser in a mental process.
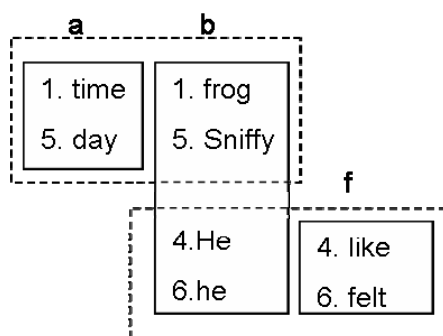

Figure 1: Chain interaction of the *frog* text

Furthermore, Hasan (1984, p. 218) proposes that examining chain interaction as such is useful for quantitatively analyzing the cohesive qualities of texts. She introduces the notion of *cohesive harmony* as a way to judge the degree of text coherence. Hasan's conception of cohesive harmony comprises the following claims:

1. The measure of cohesive harmony lies in the percentage of *central tokens*, i.e. lexical tokens which enter in chain interaction, e.g. (1) *frog*, (4) *like*, and (6) *felt* in the above text; any text will be seen as coherent if it has more than fifty percent central tokens.

2. The degree of cohesive harmony should dovetail with the degree of texts on a cline of coherence as judged by informants.

3. If two texts have similar degrees of cohesive harmony, variation in coherence will correlate with the number of breaks in the chain of interaction.

In sum, Hasan's proposals point out that examining cohesive chains, chain interaction and cohesive harmony is a useful means for demonstrating how components in a language text are deployed to construct a coherent text. To date the application of Hasan's approach to examining text coherence is restricted to language texts (cf. Parsons, 1991; Hedberg & Fink, 1996; Taboada, 2004). In the following sections, I will turn explicitly to the case of filmic discourse and explore if multimodal text coherence, particularly how components from different modes are brought to cohere, can also be construed on the basis of Hasan's proposals.

3 Filmic cohesive chains and chain interaction

This section examines to what extent Hasan's linguistic approach can be translated to filmic text. Before broaching the topic of filmic cohesive harmony, I give a synopsis of how the two types of filmic cohesive chains, i.e. reference (identity) chains and similarity chains, are generated.

The construct of *filmic reference chains* is based on the framework of multimodal identification/phoricity, which has been developed in detail in my previous work (Tseng, 2008). Building on Martin's (1992, p. 93-157) linguistic framework, filmic identification is developed according to the three central systems: [generic/ specific], [presuming/ presenting], [comparison/ -][1]. A distinctive feature of the multimodal framework, however, lies in the **cross-modal realization** of the participants. Whether generic or specific, a participant in a film has the potential to be *presented* or *presumed* **simultaneously** in different modes. For instance, a man can be simultaneously presented in a text as a visual figure *a man* with his name written on the screen and accompanied by his theme song in the soundtrack. The

---

[1] See Tseng, (2008) for a detailed description of the system networks of filmic identification and phoricity.

presuming of a participant's identity can also be realized simultaneously in different modes. For example, after a man is presented as a visual figure in a filmic text, he can explicitly reappear throughout the filmic text as the same visual figure again or he can be mentioned in spoken texts, in written texts, and so on. Even if his visual reappearance is inexplicit (for instance, only some physical parts are revealed), as long as certain filmic resources, e.g. the strategies of continuity editing, are mobilized to function as **filmic deictics**, the inexplicit reappearance can still cue the viewer to the identity of the participant. In other words, the multimodal cohesion framework as such consists of resources from a variety of semiotic modes that can be employed to construct **cross-modal referencing** throughout a film.

Filmic similarity chains, similarly to those in language, consist of actions and qualities (and quantity) of the participants. Whether visual or verbal, the qualities and actions of participants are categorized in each similarity chain according to the lexical cohesive ties defined by Hasan (1984; 1989). Filmic similarity chains are henceforth termed *action/quality chains* in this study.

Due to space constraints, I am not able to delineate the contribution of each delicate choice of the filmic identification/phoricity systems to the creation of cohesive chains; instead I will briefly demonstrate the construct of reference and action/quality chains by exemplifying the chains in a filmic text extracted from the documentary film *Comandante*, in which the film director Oliver Stone conducts interviews with Fidel Castro over three days on various issues.

The transcription[2] of the filmic text is displayed in Figure 2. It is the beginning part of an interview where Castro narrates his memory of his contact with Nixon in 1959.

image	visualtrack	image description	verbal text
1		Colour image. Some people around the table are having lunch. Long shot.	off-screen voice (Stone): "Tell me about Nixon"
2		Colour image. Old Castro's face in very close shot	

---

[2] Each filmic image in Figure 2 is reproduced and redrawn from the film because of copyright issues; a short description is attached next to each image to compensate for the different image quality resulting from the redrawing process.

3		Black and white image in archive film. Nixon is sitting in a car and waving at people on the street. Very long shot	Off-screen voice (Castro): "The first time I met him was in March 1959."
4		Black and white image in archive film. Young Castro is getting off a plane with some other people. Medium long shot.	
5		Black and white image in archive film. Nixon is gazing at the camera. Medium close shot.	Off-screen voice (Castro): "At that time he was vice president."
6		Colour image. Old Castro is sitting in his office. Medium long shot.	
7		Colour image. Old Castro. Medium close shot.	Castro: "we talk to each other."
8		Colour image. Old Castro. Medium close shot.	Castro: "I explain our problem to him and how it possibly could be solved and he immediately drafted a report. He wrote a report"
9		Black and white image in archive film. Young Castro and Nixon are shaking hands. Medium shot.	Off-screen voice (Castro): "in which he said I was a communist."

Figure 2. A filmic text extracted from *Comandante*.

For the purpose of examining chain interaction, both types of cohesive chains as generated from the *Castro* text are displayed in Figure 3. Six reference chains (RCs) and seven quality/action chains (AQCs) can be isolated.

The reference chains in this text show that there are at least six **salient** elements which have the status of being participants: *people*, *Nixon*, *Castro*, *time, problem, report.* The principle for chain formation here is taken over from Hasan's approach. For example, a chain must contain at least two components to 'have the status' for participating in chain interaction. Thus, the reference chain of the participant *Oliver Stone* is excluded from the above formations because it comprises only one item "me" in image 1. For the same reason, the

visual actions found in the text such as [hand shaking] in image 9, [walking] in image 4, or other visual quality attributes such as *very long shot* in image 3, *very close shot* in image 2, etc. are not included in the chain formations.

(I) Reference chains

participant image	*people*	*Nixon*	*Castro*	*time*	*problem*	*report*
1	[people]	"Nixon"				
2			[Castro]			
3	[people]	[Nixon] "him"	"I"	"time, March1959"		
4	[prople]		[Castro]			
5		[Nixon] "he, vice president"		"that time"		
6			[Castro]			
7		"we"	"we", [Castro]			
8		"him, he, he"	[Castro], "I, our"		"problem, it"	"report, report"
9		[Nixon] "he"	[Castro], "I, communist"			

(II) Action/quality chains

a	b	c	d	e	f	g
1. "tell" 7. "talk", [talking] 8. "explain", [talking] 9. "said"	3. "was" 5. "was" 9. "was" 5. [existing] 6. [existing]	8. "draft" 8. "wrote"	5. MCS 7. MCS 8. MCS	1.B&W 2.B&W 6.B&W 7.B&W 8.B&W	3. colour 4. colour 5. colour 9. colour	4. MLS 6. MLS

Figure 3. Reference chains and action/quality chains of the *Castro* text. (Transcription conventions: ""= spoken text, [ ]= visual elements, MCS= medium close shot, MLS: medium long shot, B&W=black and white images)

Moreover, the reference chains in this text also show that some of the participants are presented or presumed **cross-modally**. For instance: Nixon in image 1 is first presented in the spoken text "Tell me about *Nixon*" and, in image 3, he reappears as a visual figure in the image; and Castro is presented as a visual figure in image 2 and is realized again in image 3 in his spoken text as "I". Thus, a filmic reference chain can be treated as a pointer to the nature of the cross-modal referencing because it uncovers how each participant is realized mono- or cross-modally throughout a text.

Action/quality chains in this text are comprised of three different types of actions (chains a—c) and four types of quality units (chains d—g). The AQCs *a* and *b* are also realized **cross-modally**. Chain *a* consists of verbal components such as "tell", "talk", "explain", etc. and their visual synonymy, namely, the **action** of [talking] found in image 7

8

and 8. In addition, chain *b*, which denotes the non-action 'being' process type, also consists of the verbal token "was" and the visual token [existing], with which Nixon and Castro in image 5 and 6 are seen as simply 'being there', rather than performing any actions.

Furthermore, the components in the AQCs *d–g* depict the participants' **visual quality attributes** which are produced from the different uses of camera or editing techniques. Some of these visual attributes result from the manipulation of the quality of the image in which participants appear. For instance, the different uses of lighting patterns can bring about change in a participant's appearance; and the use of black-and-white or colour films can also lead to changes in the filmic participants' physical colours. Chain *e* and chain *f* are such examples. The components in the two chains are concerned with different types of image colours: in images 3, 4, 5 and 9, the participants are depicted in black and white due to filmmaking techniques in 1959; while in images 1, 2, 6, 7 and 8, which are produced at the present time, the participants are visually realized in colour. Furthermore, some of the visual quality attributes can be manipulated by different uses of camera distances, which bring about a size change of participants in filmic image frames. For instance, the components in chain *d* and chain *g*, namely, the uses of medium close shot and medium long shot, change the physical size of Nixon and Castro: in images 5, 7, and 8, Nixon and Castro can be seen as being 'larger' than in images 3, 4, or 6 because they are placed closer to the camera.

Following this construction of the RCs and AQCs in the *Castro* text, the interaction of the chains based on Hasan's method is shown in Figure 4.
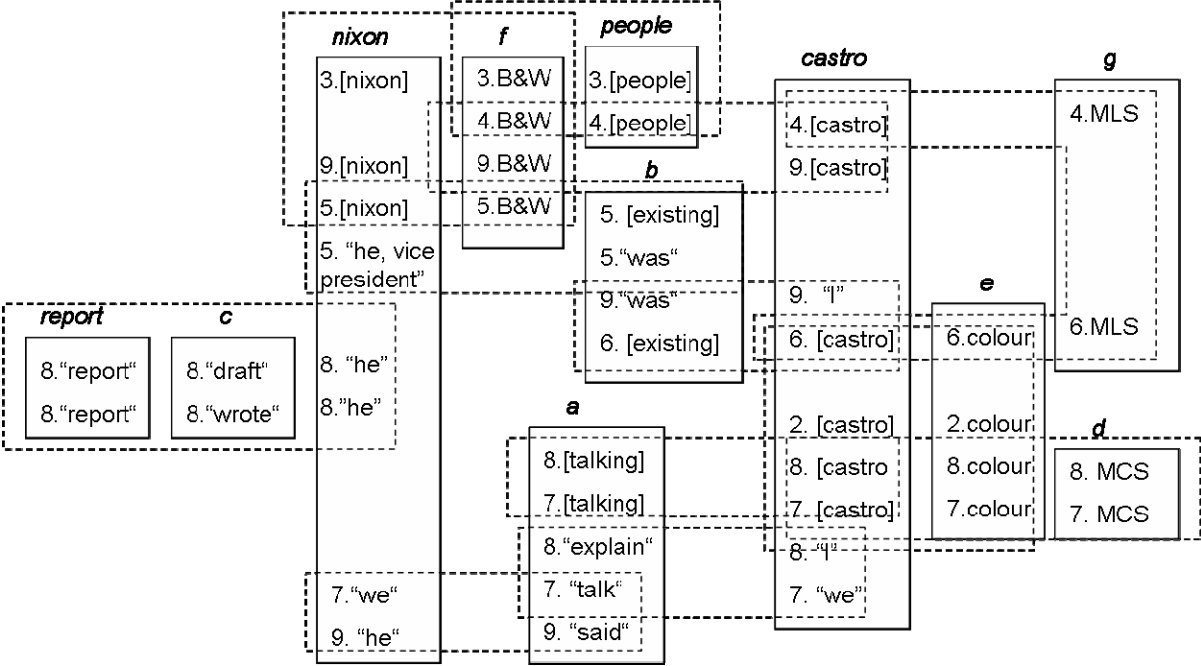

Figure 4. Chain interaction in the *Castro* text based on Hasan's method.

The chain interaction as such shows how filmic RCs and AQCs, whether mono-modal or cross-modal, are connected. However, though filmic cohesive chains can be cross-modally generated, the actual interacting components are still only linked to each other by **mono-modal ties**; that is, each individual part of a chain contributes only to a multivariate structure established within one mode. For example, in Figure 4, we can see that the verbal elements "he", "we" in the cross-modal *Nixon* chain interact only with other verbal chains such as "draft" and "wrote" in chain *c*, or with verbal elements in other cross-modal chains such as "was", "talk" in chain *a* and chain *b*. Likewise, the elements which interact with visual [Nixon] are restricted to visual actions [existing] and visual quality [B&W].

In brief, Hasan's method of chain interaction does **not** indicate **cross-modal cohesive harmony**, i.e. to what degree the units in different modes in a multimodal text are brought together to cohere in a multimodal text. Chain interaction of this kind can be used to examine the cohesive quality in verbal text and in visual images separately, and based on this method, we can also calculate the ratio of central tokens to non-central tokens. But the percentage that results only indicates the sum of two separate phenomena—visual and verbal cohesive qualities.

In order to establish multimodal text coherence, it is necessary to take into account not only the semantic ties between RCs and AQCs in each mode, but also how units in different modes are connected to establish cross-modal synergistic ties. This aspect has been partly touched upon above when we discussed the creation of cross-modal cohesive chains, which indicate how resources of different modes can be employed to realize one identity, one action, one quality and so on. But now we will take one further step and specifically address the questions as to what other types of cross-modal cohesive ties can be constructed alongside those which realize the cross-modal RCs and AQCs and if there can be a direct approach to explicitly demonstrating how the cross-modal units co-operate to achieve a coherent whole.

The different types of cohesive ties possible are displayed diagrammatically in Figure 5. We began this section with the cross-modal realization of cohesive chains, namely, how ties B1 and B2 are established, and then went on to apply Hasan's method to filmic text. The result is that Hasan's approach only indicates semantic ties of type A1 and A2. Hence, if we aim to explicitly reveal the cross-modal realization of meaning, i.e. the connections established by the synergy across different modes, not only should we take into account B1 and B2, that realize cross-modal cohesive chains, but the possible existence of the ties B3 and B4, which connect RCs of one mode to AQCs of another mode, should also be explored.
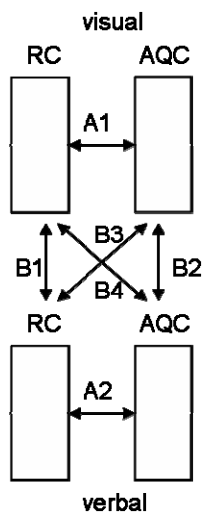
Figure 5. Establishing cross-modal interaction.

Following the elucidation of the kinds of ties to be considered in examining cross-modal coherence, the task remaining is to uncover principles for analyzing configurations of such ties. One straightforward method is to probe to what degree the chains in each mode are involved in the cross-modal synergy based on lexical cohesive relations defined by Halliday and Hasan (1976), e.g. synonymy, antonymy, meronymy, hyponymy, collocation, etc.

Cross-modal lexical cohesive bonds of this kind have also been discussed by Royce (1998) when describing intersemiotic relations in page-based multimodal text. Royce draws on Halliday's metafunctional model and formulates a framework of intersemiotic relations between the visual and verbal contributions. His work encompasses how the different semiotic modes in a printed multimodal text work together ideationally, interpersonally, and textually, and how the modes semantically complement each other to constitute a coherent whole. He coins the term *intersemiotic complementarity* to describe the synergy of the visual and verbal text. The approach to filmic cross-modal ties I am proposing here can be treated as an expansion from the ideational[3] aspect of Royce's paged-based framework into a complex method for examining synergistic ties of cross-modal units in stretches of filmic text.

To demonstrate the approach, I will exemplify the cross-modal cohesive ties found in the *Castro* text. First, the cohesive chains displayed in Figure 3 are remodelled into two groups of chains: (1) RCs and AQCs in the visualtrack and (2) RCs and AQCs in the verbal text in Figure 6. The chains which are involved in lexical cohesive relations are connected

---

[3] Hasan (1984) and Martin (1992) both points out that at present examining cohesive harmony and chain interaction is restricted to only the ideational meaning of a text. In this study it has not been possible to investigate how to examine the interpersonal aspect of cohesive analysis. But certainly in future analyses of filmic cohesion, the calibration of cohesive harmony with attitudinal/interpersonal analyses must also be conducted.

with dotted lines in the diagram. For instance, the visual *Nixon* and *Castro* chains are linked to their verbal synonymy *Nixon* and *Castro* chains. Likewise, the ties which connect the verbal *was* and *tell* chains to the visual *existing* and *talking* chains, are also realized by synonymy. These semantic bonds, shown as B1 and B2 in Figure 5, are precisely the ties which realize the creation of the cross-modal cohesive chains elucidated earlier in the section.

chains in the visualtrack

people	Nixon	Castro	MCS	talking	exsiting	B&W	colour	MLS
1,2,4.	3,5,9	2,4,6, 7,8,9	5,7,8	7,8	5,6	1,2,6, 7,8	3,4, 5,9	4, 6

syno. syno. collo. syno. syno.

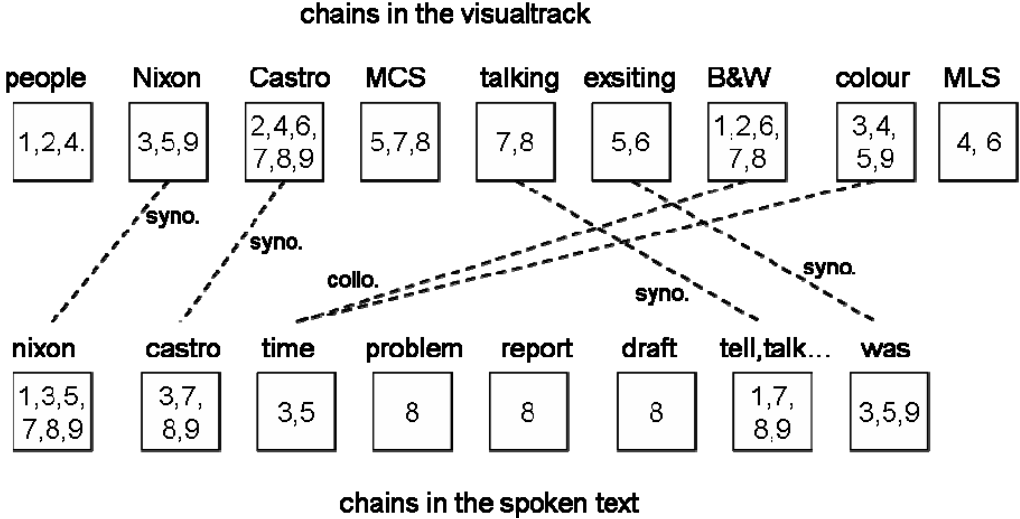| nixon	castro	time	problem	report	draft	tell,talk…	was
1,3,5, 7,8,9	3,7, 8,9	3,5	8	8	8	1,7, 8,9	3,5,9

chains in the spoken text

Figure 6. Cross-modal interaction in the *Castro* text (the numbers in chains: indicating images where components are realized).

Furthermore, the chains of *colour* and *B&W* in the visualtrack are now seen interacting with the verbal *time* chain—whether a participant is a colour or black-and-white visual figure is mostly related to and co-occurs with the configurations of historical time (e.g. 1959 or the present). Therefore, the cohesive ties which connect the *colour* and *B&W* chains to the *time* chain are established by the relation of collocation. The cohesive tie of collocation is precisely what Hasan excludes in the formation of cohesive chains and chain interaction (Hasan, 1984). However, such types of ties (shown as B3 and B4 in Figure 5) are in fact not uncommon in filmmaking and would be overlooked if we merely examine the formation of cohesive chains for the phenomenon of cross-modal coherence. Moreover, establishing such cohesive ties (connecting an AQC and a RC across different modes) is significant in construing symbolic meanings in certain types of film, for instance, films that employ the particular strategies of *montage* developed by the Soviet filmmakers of the 1920s such as Pudovkin and Eisenstein (cf. Eisenstein, 1969). I will exemplify this point by analysing the semantic ties of a sequence in Eisenstein's film *Strike* on the basis of Vogel's (1974) synopsis of how Eisenstein's montage worked in practice:

"The best example of Eisenstein's montage methods occurs in the famous sequence in which the four capitalists dealing with the strike are seated in the plush comfort and isolation of their mansion, smoking and drinking. Through cross-cutting, we now see, in this order: workers at a clandestine strike meeting; the capitalist putting a lemon into a juice extractor; the workers discuss their demands; the handle of the juice extractor descends to crush the fruit; the workers are charged by mounted police; the boss says in an intertitle: 'Crush hard and then squeeze!'; the workers are attacked; a piece of lemon drops on the well polished shoe of the capitalist; disgusted, he uses the paper containing the workers' demands to wipe it off". (p. 36)

Some multivariate and covariate ties linking the message units in this sequence described by Vogel are shown in Figure 7. In the visualtrack, though the scenes of police attacking workers and capitalist squeezing lemon with juice extractor are presented as two discontinuous, unrelated events, the visual action chain of *squeeze/attack*, which conjoins the two events, explicitly lays bare their metaphorical association created by montage. Moreover, according to Vogel's description, the verbal text in the intertitle is actually edited between the scenes of workers being charged by police and workers being attacked. That is to say, the intertitle is meant to refer to the command of the capitalist boss to 'crush the workers harder' because the viewer indeed sees the workers being 'attacked harder' right after the intertitle. Nevertheless, the cross-modal cohesive ties displayed in Figure 7, i.e. the relation of collocation established between the verbal AQC the *crush/squeeze* chain and the visual RCs, the *juice extractor* and *lemon* chains, further highlight the metaphorical bonds of crushing lemon and attacking workers commanded by the capitalist.
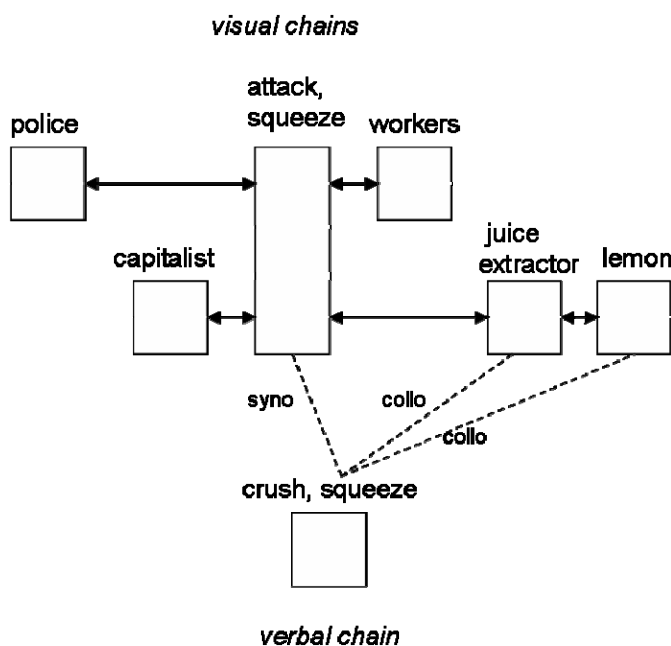

Figure 7. Semantic ties linking units within and across modes in *Strike* (arrows=mono-modal multivariate ties, dotted lines=cross-modal covariate ties).

13

Establishing filmic semantic bonds in this way gives a clear picture of how units of different modes are brought together to cohere; and it can also be considered as an important starting point, not only for developing a more complex approach to quantitatively analysing the ratio of filmic cohesive harmony of different filmic texts, but also for exploring the correlation between multimodal text coherence and filmic genre. In the next section, I will suggest some possible extensions of the approach proposed in this paper for the further formulations of tools for investigating texts structure and filmic genre.

4 Cohesion, coherence, and further developments of filmic analysis

As a working hypothesis, I propose that filmic cohesive chains and cohesive harmony can do at least two things with regard to the further development of filmic discourse analysis:

First, filmic cohesive chains can be used to reflect generic stages of filmic texts. So far this finding has been restricted to language genres (cf. Hasan, 1984; Ventola, 1987; Taboada, 2004). Nevertheless it is suggested that it can be expanded to filmic text along the same lines as the linguistic approach. Take the structural movement in the *Castro* text for example. The reference chains displayed in Figure 3 explicitly track the two 'protagonist participants' *Nixon* and *Castro* throughout the text. In addition, the participation of new chains or the removal of old chains also launches the start of another stage. For instance, the removal of the *people* and *time* chain symbolizes the ending of the 'orientation stage' which depicts the background information of the contact between Nixon and Castro, and the participation of *problem* and *report* chains opens up another unit in which more information regarding Nixon and Castro's meeting is to be given.

Second, in filmic texts, the variation of cross-modal cohesive ties might also reflect certain aspect of filmic genre. Van Leeuwen's (2005) analysis of news reading in part bears out this point. By demonstrating how rather dramatic news content on a strike at an airport is visually accompanied by a comparatively 'quiet' scene in the airport, his analysis points to the lower degree of cross-modal cohesive harmony commonly seen in the genre of news reports. In van Leeuwen's words:

> "In television news items the image track often hangs together like loose sand and would not make much sense on its own. Here the key principle behind the editing is spatial co-presence: all the shots have been taken in or around the airport. There is only the occasional sense of 'next event' and 'detail/overview' linking… The images merely serve as a kind of setting, which gives a spurious sense of authenticity to the news item. Spurious, because, although this item is about a strike, we do not see the workers striking, in fact we see them

loading suitcases into a plane. Not the events are illustrated here, but only actors…and location… This is by no means an exception". (p. 237)

From van Leeuwen's analysis, we can draw an assumption that cross-modal cohesive harmony is relatively low in the genre of news reports because they have lower degrees of cross-modal chain interaction—in most cases, only reference chains (settings, actors) are involved in the cohesive ties while action/quality chains in visual and verbal modes are unrelated.

While the correlation of text coherence and genre (see Hasan, 1989, for a detailed view of the relations between genre and texture in language) has been widely discussed in language texts (cf. Parsons, 1991; Hedberg & Fink, 1996; Taboada, 2004), in multimodal text, more methods for investigating cross-modal ties, e.g. interpersonal synergistic relations, are still needed to provide a substantial insight into the correlation between filmic genre and the configuration of filmic texture.

5 Some brief conclusions

I began this paper by advocating the need to develop more 'bridging devices' for the analysis of cross-modal meaning realization. As one step towards this goal, I went on to apply Hasan's method concerning coherence and cohesive harmony to filmic texts. It was found that chain interaction according to Hasan's principle does not indicate cross-modal interaction. I proposed a method for examining how units in each mode interact to establish a coherent whole. This approach takes into account not only the cohesive ties within cross-modal cohesive chains, but also the ties which move beyond the boundary of cohesive chains and link identity of one mode to action or quality of another mode. Further, I also suggested some possible implementations of filmic cohesive chains and the cross-modal approach proposed in this chapter with regard to the further development of tools for filmic discourse analysis.

More issues can be raised following the application of Hasan's approach to film. We can ask questions such as can there be an approach for quantitatively analysing and comparing degrees of cohesive harmony in different filmic texts? Or in what way does filmic genre act upon the correlation between cross-modal and mono-modal cohesive harmony in a text? For instance, does a higher degree of cross-modal cohesive harmony with lower degree of mono-modal cohesive harmony, or high degree in both, also reflect a distinctive aspect of filmic genre?

Furthermore, by the start of the 21st century, in English speaking cultures at least, film had become an increasingly significant aspect of the school curriculum, though teachers

are not always very skilled in teaching about it. Publications have begun to appear offering practical advice to teachers for media study generally, including film (e.g. Callow, 1999; Unsworth, 2001); nevertheless, more work concerning filmic analysis as part of curriculum remains to be done in the future.

Finally, through the development of filmic cohesion and coherence, I hope to have widened the possibilities of the further formulation of an extensive filmic model which permits the investigation and comparison of filmic texts across types and genres; and this goal is precisely what film theorists over decades have been striving to achieve.

References

Baldry, A. P.& Thibault, P. (2006). *Multimodal transcription and text analysis. A multimedia toolkit and coursebook*. London: Equinox.

Bateman, J. A. (*forthcoming*-a). *Multimodality and genre*. London: Palgrave Macmillan.

Bateman, J. A. (*forthcoming*-b). Towards a 'Grande Paradigmatique' of film: Christian Metz reloaded. Paper accepted for publication in *Semiotica* .

Bordwell, D. (1984). *Narration in the fiction film*. Madison: University of Wisconsin Press.

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

Callow, J. (ed.). (1999). *Image matters: Visual texts in the classroom.* Sydney: Primary English Teaching Association.

Eisenstein, S. (1969). *Film form: Essays in film theory*. Harvest Book.

Gledhill, C. (1985). Genre. In Cook P. & Bernink M. (ed.), *The cinema book*. London: British Film Institute.

Halliday, M. A.K. & Hasan, R. (1976). *Cohesion in English*. London: Longman.

Hasan, R. (1984). Coherence and cohesive harmony. In Flood J. (ed.), *Understanding reading comprehension:cognition, language and the structure of prose*. Newark, Delaware: International Reading Association.

Hasan, R. (1989). The texture of a text. In Halliday M. A. K. & Hasan R. (eds.), *Language, context and text: Aspects of language in a social-semiotic perspective*. Oxford University Press.

Hedberg, N. L. & Fink, R. J. (1996). Cohesive harmony in the written stories of elementary children. *Reading and Writing* 8(1), 73–86.

Kress, G. & van Leeuwen, T. (2001). *Multimodal discourse: The modes and media of contemporary communication*. London: Arnolds.

Macken-Horarik, M. (2004). Interacting with multimodal text: reflections on image and verbiage in ArtExpress. *Visual Communication* 3(1): 5-26.

Martin, J. R. & Rose D. (2003). *Working with discourse*. London: Continuum.

Martin, J. R. (1992). *English text: System and structure*. Amsterdam: Benjamins.

Mittell, J. (2001). A Cultural approach to television genre theory. *Cinema Journal* 40(3), 3–24.

Mittell, J. (2004). *Genre and television*. New York and London: Routledge.

Neale, S. (2000). *Genre and Hollywood*. London: Routledge.

Nichols, B. (2001). *Introduction to documentary*. Bloomington: Indiana University Press.

O'Halloran, K. (ed.). (2004a). Multimodal discourse analysis: Systemic functional perspectives. London: Continuum.

O'Halloran, K. (2004b). Visual semiosis in film. In O'Halloran K. (ed.), *Multimodal discourse analysis: Systemic functional perspectives*. London: Continuum.

Parsons, G. (1991). Cohesion coherence: Scientific texts. In Ventola E. (ed.), *Functional and systemic Linguistics*, Berlin: Mouton de Gruyter.

Royce, T. (1998). Synergy on the page: Exploring intersemiotic complementarity in page-based multimodal text. *JASFL Occasional Papers* 1(1), 25–48.

Taboada, M. T. (2004). *Building cohesion and coherence: Task oriented dialogue in English and Spanish*. Amsterdam: John Benjamins.

Tseng, C. (2008). *Cohesion in multimodal discourse*. Ph. D.thesis, University of Bremen, Germany.

Tudor, A. (1973). Genre. In Grant, B. K. (ed.), *Film genre reader*. Austin: University of Texas Press.

Unsworth, L. (2001). *Teaching multiliteracies across the curriculum*. Buckingham and Philadelphia: Open University Press.

van Leeuwen, T. (1985). Rhythmic structure of the film text. In van Dijk, T. (ed.), *Discourse and communication: New approaches to the analysis of mass media discourse and communication* (pp. 216–232). Berlin: Walter de Gruyter.

van Leeuwen, T. (1991). Conjunctive structure in documentary film and television. *Continuum* 5(1), 76–115.

van Leeuwen, T. (1996). Moving English: The visual language of film. In Goodman, S. & Graddol, D. (eds.), *Redesigning English - New Texts, New Identities* (pp. 81–105). London: Routledge.

van Leeuwen, T. (2005). *Introducing social semiotics*. London: Routledge.

Ventola, E. (1987). *The structure of social interaction: A systematic approach to the semiotics of service encounters*. London: Printer.

Ventola, E., Charles, C. & Kaltenbacher, M. (ed.). (2004). *Perspectives on Multimodality*. Amsterdam: John Benjamin.

Vogel, A. (1974). *Film as subversive art*. London: Random House.