Compositional meanings in Chinese subtitle translation: A multimodal analysis of Mulan

Yuping Chen
China Agricultural University, China
Wei Wang
The University of Sydney, Australia

While roles of semiotic interplay in subtitle translation have been widely acknowledged, scant research has been done to examine how visual and verbal modes really interact with each other to produce a coherent and cohesive whole of a subtitled film. Drawing on the concepts of compositional meanings and semiotic cohesion informed by systemic functional linguistics, this paper conducts a multimodal analysis of the Chinese subtitles in an English language film Mulan, focusing on examining compositional meanings from three aspects: types, functions of semiotic interrelations and impacts on linguistic parts of subtitles. This paper finds that co-referentiality (Halliday and Hasan 1985) is the most predominant cohesive relation between subtitle and image, which exerts more impacts on linguistic parts relating to identification of subtitles. This finding can be seen in two implicit encoding devices of anaphora and cataphora. It suggests that a subtitling unit is not limited to a filmic unit of shot, but goes far beyond it, which contributes to maintaining the narrative flow in subtitled films. These findings deserve to be further explored in examination of subtitling practice and construction of theoretical models in this growing field of research.

Keywords: compositional meanings, multimodal analysis, semiotic cohesion, cataphora, anaphora

1. Introduction

In recent years, the analysis of screen discourse has been much discussed in the field of audiovisual translation with a particular focus on subtitle translation. Though the multimodal feature of subtitling has been recognized in previous studies (e.g. Chuang

2006; Taylor 2003, 2004; Zhang 2012), scant research has been done to address metafunctions in subtitle translation from a multimodality perspective. Concurring with the view that examination of metafunctions is of great significance in the subtitling process, this paper aims to investigate one of these meanings, i.e. compositional meanings, through analysing the Chinese subtitles in an English language film Mulan. This film was produced by Walt Disney Pictures and released with its Chinese subtitles on 19 June 1998 in mainland China. It enjoys great popularity among Chinese viewers with its strong Chinese elements. As one of the most successful imported English language blockbusters in mainland China, the examination of Chinese subtitles in Mulan is expected to reveal patterns and regularities in subtitle translation of other similar films in mainland China.

To that end, three research questions are formed to address the following issues:

Q1: how does the visual mode interrelate with the written verbal mode (i.e. subtitles) to generate compositional meanings in the subtitled film?

Q2: how does the intermodal interplay maintain the narrative flow of the subtitled film?

Q3: what concrete influences are brought about to linguistic parts of subtitles through intermodal interplay?

This paper begins with an introduction of the theoretical background with an aim of constructing an analytical framework. Then, it moves on to the data analysis of Mulan. This paper ends with a discussion and a conclusion of this multimodal analysis in relation to the previous research.

2. Theoretical background

2.1. Compositional meanings

Halliday's Systemic Functional Linguistics (SFL) identifies three metafunctions to convey meanings embodied in human languages, claiming these metafunctions or meanings constitute the basis of the semantic organization of all natural languages. He terms these as ideational, interpersonal and textual metafunctions. To examine the meanings of images, Kress and van Leeuwen (2006) adopt these three metafunctions with three different terms: representational, interactive and compositional meanings.

Representation refers to informational content about something that is represented visually. When visual information engages either directly or indirectly with viewers, it is labelled interaction, whereas information in images, organized in line with compositional criteria, is called composition. The latter (compositional meanings) is the research focus of this paper.

Given this paper's focus on visual-verbal intercommunication, Kress and van Leeuwen's terminology "compositional" is adopted to examine how semiotic interplay functions to produce a coherent and cohesive whole of a subtitled film, which closely relates to the second key concept (i.e. semiotic cohesion) that is deployed in this paper.

2.2. Semiotic cohesion

Degree of connectedness between text and image relates to the ways in which visual mode and verbal mode correlate with each other. These semiotic connections are established through semantic ties, which form the basis of cohesion (Caple 2013). Within a verbal text, semantic ties can be realized in a number of ways, grammatically and lexically.

Halliday and Hasan (1985) introduce three of the main cohesive devices contributing to semantic ties: co-referentiality, co-classification and co-extension. Co-referentiality "is typically realised by the devices of reference, such as the pronominals 'he', 'she', 'it', etc. or by the use of the definite article 'the' or that of the demonstratives 'this' or 'that'"(Halliday and Hasan 1985:74). "Co-classification is normally realised either by substitution or by ellipsis" (Halliday and Hasan 1985: 74). As for the relation of co-extension, in the linguistic field, four concepts are adopted to form the sense relations, namely, synonym (i.e. "the experiential meaning of the two lexical items is identical"), antonym (i.e. "the oppositeness of experiential meaning"), hyponymy (i.e. "a relation that holds between a general class and its sub-classes") and meronymy (i.e. "refers to a part-whole relation") (Halliday and Hasan 1985: 80-81).

It can be seen that these cohesive devices are essentially relational, whose interpretation has to be found by reference to some other explicit sources, which can either precede or follow the implicit messages. When implicit term follows its explicit linguistic reference, this cohesive tie is anaphoric, while when the implicit term precedes its explicit linguistic referent, this cohesive tie is known as cataphoric (Halliday and Hasan 1976).

In a verbal text, cohesion focuses on the "grammatical" aspects at the sentence level

through substitution, reference, lexical chains and conjunction (Halliday and Hasan 1976). This paper aims to find out how verbal and visual information correlate with each other to create cohesion and what cohesive devices are applicable in this process.

In short, semiotic cohesion contributes to compositional meanings; cohesive devices build up semantic ties and semantic ties promote semiotic cohesion. Thus, the key to examine compositional meanings is to analyse semantic ties between word and image, which closely correlates with the building up of the analytical framework for this paper.

2.3. An analytical framework

Semiotic cohesion in this paper particularly refers to how visual-verbal interplay leads to cohesive ties. Three issues need to be addressed for the sake of constructing an analytical framework. First, which mode bears implicit information, and which mode is endowed explicit information. Second, whether cohesive relations between semiotic sources, such as word and image, are as same as those between word and word. Third, how to establish cohesive relations between word and image.

With regard to the first issue, Barthes (1984) purports that verbal elements reign over the visual counterparts. Relationships between verbal and other semiotic modes of communication are hierarchical and asymmetric. Barthes (1984) argues that verbal text is the primary code upon which other semiotic codes, including images, are dependent. Visual signs constitute a "floating chain of signifieds" that needs to be "fixed" or constrained by language (Barthes 1984: 39). Drawing on Barthes's study, visual information tends to be implicit, while verbal mode provides explicit reference to its visual counterpart. However, in the case of subtitle translation, due to the spatial and temporal constraints, subtitles are expected to be as concise as possible. Visual information is a blessing in audiovisual texts, functioning to further define or fix verbal information in most cases. This is opposite to the situation in other multimedia texts, such as newspaper or magazine, in which there are always more spaces for words than images. Therefore, this paper purports that pictorial elements provide explicit information to define verbal languages. In this vein, anaphora occurs when visual information precedes verbal expressions, while cataphora takes place when images follow words.

Furthermore, analytical units in audiovisual texts are distinctive, which are essential for anaphora and cataphora. These analytical units are the four basic filmic units: shot, scene, sequence and stage. According to Iedema (2001), in a shot the camera movement is unedited (uncut); there are more than one shot in a scene, in which all shots share

the continuity of time and space; a range of contiguous scenes enjoying a thematic or logical continuity lead to sequences; stages are those elements which tell us where we are in the overall filmic text.

Though these deictic references, pointing forward or backward, can also be found in solo-modal texts, such as print texts, and other non-filmic multi-modal texts, such as caricatures, it has not been fully applied in subtitle translation research. Possibly this is due to the fact that the edited visual images in films always flash by fast, downgrading their importance in subtitle translation.

With regards to the second and the third issues, Caple argues that it is very difficult to establish cohesive relations between word and image because "photographs and the words that accompany them can never enjoy a one-to-one relationship" (2013: 144). So, the basic units of analysis need to be identified for multimodal analysis. Drawing on Caple (2013) and Royce (2002), the ideational features of identification, activity, circumstances and attributes will be employed in this paper to compare cohesive relations between word and image.

- *identification: Who or what are the represented participants, or who or what is in the visual frame (animate or inanimate)?*
- *activity: What processes are there, or what action is taking place between the actor(s) and the recipient(s) or object(s) of that action?*
- *circumstances: What are the elements that are locative (i.e. concerned with the setting), are of accompaniment (i.e. participants not involved with the action), or are of means (i.e. participants used by the actors)?*
- *attributes: relating to the participants' qualities and characteristics*
  (Royce 2002: 193−194)

Thanks to these four features, visual and verbal information are unified and become comparable and identifiable to each other.

Intermodal cohesive relations between ST (source spoken verbal text) and images (hereafter Semiotic system I), and that between TT (target written verbal text, i.e. subtitles) and images (hereafter Semiotic system II), are examined respectively with a view to revealing whether different cohesive relations are preferred in these two systems. Also, which linguistic part of subtitles is most heavily influenced in subtitling is investigated by relating to whether it is identification (nouns, pronouns and nominal phrases), activity (verbs and phrasal verbs), circumstances (nouns and prepositional phrases) or attributes (adjectives, adverbs, adjective phrases and adverbial phrases) is affected. Furthermore, whether the transformations of cohesive relations between Semiotic

system I and Semiotic system II promote the narrative flow is investigated. To maintain the narrative flow is one of the key functions of subtitle translation (Chen and Wang 2016), which guarantees coherence at the contextual level. Coherence is like the "glue" that joins ideas together to form a cohesive whole. In a verbal text, coherence mainly refers to the "rhetorical" aspects of a text, including developing, supporting, organizing and clarifying ideas. This paper tends to find out whether and how intermodal interplay promotes coherence in the subtitling process through maintaining the narrative flow of subtitled films.

An analytical framework for this study is thus constructed as shown in Figure 1.
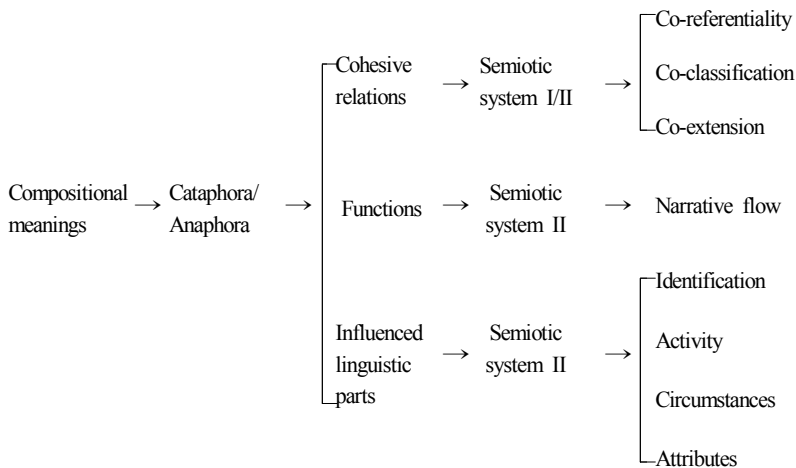

Figure 1. An analytical framework for compositional meanings in subtitle translation

Figure 1 illustrates that compositional meanings in subtitle translation are examined by resorting to two implicit encoding devices: cataphora and anaphora. Three issues are addressed: cohesive relations between visual and verbal modes (spoken verbal mode in Semiotic system I and written verbal mode in Semiotic system II); functions of semiotic cohesion in delivering compositional meanings; influenced linguistic parts of subtitles with reference to four features of ideational metafunction.

This framework is employed to analyse the Chinese subtitles in Mulan.

3. Compositional meanings in this film

Compositional meanings derived from cataphora result from intermodal relations between what is present visually in subsequent filmic units and current subtitles, while in anaphora they are derived from intermodal interplay between what has already been present visually and current subtitles. Based on the data of Mulan, it is found that there are more instances relating to anaphoric device than to cataphoric device. See Table 1 below.

Table 1. The numeric summary of the instances relating to cataphora and anaphora

Cataphora	Anaphora
9	24

Table 1 shows that there are more instances involving semiotic interplay in anaphora (24 instances) than in cataphora (9 instances), indicating that information visually expressed information in preceding filmic units is more involved in delivering compositional meanings in the subtitling process than pictorial information in subsequent filmic units.

3.1. Compositional meanings driven by cataphora

Roles and functions of semiotic cohesion and what linguistic parts of subtitles are influenced are examined.

3.1.1. Roles and functions of semiotic cohesion

The data below shows that cohesive relations in Semiotic system I have undergone changes when they are transferred to Semiotic system II. Examination of these changes reveals how visual mode interrelates with verbal mode to generate compositional meanings in subtitle translation (see Table 2 below).

Table 2. Cohesive ties between text and image based on cataphora

Ideational features	Ideatio	ST	Cohesive ties I	TT	Cohesive ties II
1.Identification	1.Huns	1. -	1.C*	1.匈奴入侵	1.R*
2.Identification	2.signal fire	2.the signal	2.C*	2.烽火	2.R*
3.Activity	3.pour water over	3.Put it out	3.E*	3.浇水	3.R*
4.Attributes	4.collapse of a statue	4.Oh!	4.E*	4.惨了	4.R*
5.Activity	5.play drum	5.Take it!	5.E*	5.敲吧	5.R*
6.Activity	6.dress up	6.to work with	6.E*	6.装扮	6.R*
7.Identification	7. notes	7.Notes	7.C*	7.小抄	7.R*
8.Activity	8.feed chicken	8.my chore	8.C*	8.喂鸡	8.R*
9.Activity	9.be watched	9.show 'em	9.R	9. -	9.C

Note*: C stands for co-classification; R stands for co-referentiality; E stands for co-extension

Table 2 displays that when cataphora is concerned, the most frequently demonstrated role of visual mode in Semiotic system I (i.e. between ST and images) is co-extension (4 out of 9 instances) and co-classification (4 out of 9 instances) (as shown in Cohesive ties I), while the most predominant role of image in relation to TT (i.e. subtitles) is co-referentiality (8 out of 9 instances) (as shown in Cohesive ties II). What is deserved to note is that all co-classification ties and co-extension ties in Semiotic system I change to co-referentiality tie in Semiotic system II.

Instance 1 is provided below, showing how co-classification cohesive relation in Semiotic system I is transformed into co-referentiality relation in Semiotic system II. Meanwhile, the function of such a change is interpreted as well.

Instance 1: [Scene at 01:37-02:03]
ST: *We are under attack! Light the signal!*
TT: *匈奴入侵，快点烽火！*
    (Huns are intruding! Light the signal fire!)

The subtitle *匈奴入侵，快点烽火("Huns are intruding! Light the signal fire!")* is present in the shot at 01:37, in which only a visual image of a soldier shouting "We are under attack! Light the signal" is shown. In the subsequent shots at 01:42 and at 02:03, the images of Huns and the signal fire on the Great Wall are present on screen. ST does not provide the specific agent of the action "attack" and the concrete

connotation of "signal", but the explicit meanings of these two implicit terms are provided visually. Since these two pairs of implicit and explicit terms are members of two classes (agent of an action and Huns; signal and the signal fire on the Great Wall), the interplay between ST in the shot at 01:37 and the images at 01:42 and at 02:03 establishes a co-classification tie in Semiotic system I, which changes to a co-referentiality tie between TT and images in Semiotic system II, when "signal" is rendered into 烽火("the signal fire on the Great Wall") and when 匈奴("Huns") is added in the subtitle.

Though the literal translation 我们受到了攻击("We are under attack") and 有人攻击 ("Under attack") can also transfer the meaning of ST, they fail to visually relate this shot to the following turn, thus weakening the narrative flow. "Each character intervention within the dialogue constitutes one turn" (Remale 2003: 229, italicized in original). In film, a dialogue contributes to making up one turn by modifying the previous turn(s) and anticipating the next, which "potentially modifying the entire narrative situation" (Remale 2003: 229). The narrative flow is thus maintained when

- participants are presented on screen and verbally simultaneously in the same filmic unit;
- participants are presented on screen and verbally in different filmic units;
- participants are presented only on screen or only verbally in one filmic unit only.

In this paper, all instances showing co-referentiality tie between TT and images demonstrate the second type of narration flow. Then the question arises: why is it necessary to present both verbally and by image in subsequent filmic units to strengthen the narrative flow? This paper argues that if this redundant presentation does not go against the need to overcome the technical constraints in subtitling, there would be no reason to oppose to do so.

Two more similar instances are provided below, showing how visual-verbal presentation in different filmic units help maintain the narrative flow. Instance 2 reflects the transformation of cohesive tie from co-classification to co-referentiality, and Instance 3 demonstrates the change of cohesive tie from co-extension to co-referentiality.

Instance 2: [Sequence at 06:45-10:05]

ST: ···Notes···

TT: ···小抄···

    (···A piece of cheat sheet···)

The translation of "Notes" into 小抄 ("a piece of cheat sheet") in the shot at 06:45 is justified by the image at 10:05, in which it is visually shown that Mulan peeps at the words written on her forearm when she was questioned a question by a matchmaker. The introduction of the visual participant (a piece of cheat sheet) in the subsequent filmic shot helps the maintaining of the narrative flow because of the change of cohesive tie from co-classification (notes and a piece of cheat sheet belong to the same class of written words) in Semiotic system I to co-referentiality (小抄 identifies "a piece of cheat sheet") in Semiotic system II.

> Instance 3: [Scene at 23:56-24:03]
> ST: *Oh*!
> TT: 惨了!
>   *(So awful!)*

The source verbal information "Oh" is translated into 惨了 ("So awful!") in the shot at 23:56. This is because of the visual image shown in the shot at 24:03, in which a giant stone dragon collapsed to the ground because of Mushu's (i.e. a little dragon) mischief. The cohesive tie of co-extension ("Oh" extends to "miserable") thus changes to co-referentiality (惨了 equals "So awful"). The co-presentation of verbal and visual information in these two consecutive shots facilitates the maintaining of the narrative flow by anticipating the visual content in the next turn.

The reason why overcoming the technical constraints is not an essential function when semiotic interplay delivers compositional meanings is that when subtitle translation is carried out in filmic units beyond shots, long subtitles can be split into several parts and be present in two or even three consecutive filmic units. Consequently, there are no spatial or temporal technical constraints to overcome.

The above-mentioned is the analysis of roles and functions of semiotic cohesion when cataphora is applied in delivering compositional meanings in subtitle translation. With a view to exploring to what extent semiotic cohesion can be demonstrated linguistically in subtitles, the next sub-section concerns how linguistic parts of subtitles are influenced because of semiotic cohesion.

3.1.2. Influenced linguistic parts

It can be seen from Table 2 above that the most heavily influenced linguistic part is that relating to identification, which includes most nouns, pronouns or nominal phrases, and that concerning activity, which involves most verbs and phrasal verbs, while the

least influenced part is that correlating with attributes, most of which are modifiers of participants qualities, such as adjectives (phrases) and prepositional phrases. No instances involving the expression of circumstances is found. Instance 4 below demonstrates how linguistic parts relating to identification and activity are affected in the subtitling process.

> Instance 4: [Scene at 06:11-06:56]
> ST: *This is what you give me to work with.*
> TT: *这是我要装扮的女孩吗?*
>   *(Is this the girl I am going to help to dress up?)*

The visual presence of Mulan in the shot at 06:11 justifies the added message *女孩* ("the girl") in the subtitle, and the visual image of dressing up shown at 06:56 further explains the translation of "work with" into *装扮* ("to dress up") in the subtitle at 06:11. The visual information of the participant (i.e. Mulan) and the activity (i.e. dressing up) is incorporated in the subtitle. The literal translation *这就是你们给我要弄的* ("This is what you give me to work with"), on the one hand, sounds more vulgar, and on the other hand, this implicit Chinese version forces viewers to refer to both the visual mode and the written verbal mode to ensure the complete comprehension of the message, which imposes more processing burden on viewers. This is because "[s]ubtitled films likely tax the attention and memory systems because there is visual information ⋯ as well as verbal information ⋯ one must switch from subtitles to visual scene and vice versa to understand the story" (Lee et al. 2013: 414).

The above analysis concentrates on how visual information present in subsequent filmic units influences the translation of subtitles. In addition to this situation, there are also cases where visual messages presented previously exert impacts on subtitle translation. This is how compositional meanings are delivered because of the implicit encoding device of anaphora.

3.2. Compositional meanings driven by anaphora

Roles and functions of semiotic cohesion and what linguistic parts of subtitles are influenced are examined as well.

3.2.1. Roles and functions of semiotic cohesion
Three types of cohesive relations between visual and verbal modes, i.e.

co-referentiality, co-classification and co-extension, are also examined in Semiotic system I and Semiotic system II. Comparison and contrast of how these cohesive relations move between these two systems discloses how visual mode interrelates with verbal mode to deliver compositional meanings in subtitle translation. See Table 3 below.

Table 3. Cohesive ties between text and image based on anaphora

Ideational features	Image	ST	Cohesive ties I	TT	Cohesive ties II
1.Identification	1.a dragon	1.serpentine salvation	1.C*	1.匈奴入侵	1.R*
2.Attributes	2.stone on the body	2. -	2.E*	2.烽火	2.R*
3.Attributes	3.happy	3.Yep	3.E*	3.浇水	3.R*
4.Identification	4.an old man	4. -	4.R*	4.惨了	4.R*
5.Identification	5.fire dies out	5.pretty hot	5.E*	5.敲吧	5.R*
6.Attributes	6.messy	6.impressive	6.C*	6.装扮	6.R*
7.Attributes	7. timid	7.some	7.E*	7.小抄	7.R*
8.Identification	8.a military letter	8.the army	8.C*	8.喂鸡	8.R*
9.Activity	9.hug	9.do that	9.E*	9. -	9.R*
10.Identification	10.stone dragon statue	10.Stone Humpty-Dumpty	10.C*		10.R*
11.Identification	11.tattoo	11. -	11.E*		11.R*
12.Activity	12.slap	12.slap	12.R*		12.C*
13.Identification	13.a silly boy	13.lunatic	13.R*		13.R*
14.Activity	14.bath	14.hygiene	14.E*		14. -
15.Circumstances	15.Imperial city	15.city	15.C*		15.R*
16.Identification	16.rice	16.a single grain of rice	16.R*		16.E*
17.Identification	17.a panda courier	17.a black-and-white	17.C*		17.C*
18.Activity	18.drill	18.work	18.E*		18.R*
19.Activity	19.bite the butt	19.owe me big	19.E*		19.R*
20.Attributes	20.disdain	20.They don't even like me	20.E*		20. -
21.Activity	21.safe and sound	21.like daisies	21.E*		21.R*
22.Identification	22.blow off the mountain	22.the mountains	22.E*		22.R*
23.Identification	23.Great Wall		23.C*		23.R*

	24.cross-dress as		
24.Activity	a man to join the	24.E*	24.R*
	army		

Table 3 shows that co-extension (13 out of 24 instances) is the most dominant cohesive tie in Semiotic system I and the least used cohesive tie is co-referentiality (4 instances), while the most frequently demonstrated cohesive tie in Semiotic system II is co-referentiality (19 out of 24 instances). Furthermore, all those instances showing co-extension tie in Semiotic system I (as shown in Cohesive ties I) are transformed to co-referentiality relation in Semiotic system II (as shown in Cohesive ties II).

Instance 5 below illustrates how co-extension relation is transformed to co-referentiality, while Instance 6 shows how co-classification relation changes to co-referentiality. Also, how the narrative flow is better maintained because of these changes is interpreted.

Instance 5: [Scene at 1:17:59-1:18:03]

ST: *Is she allowed to <u>do that</u>?*

TT: *她竟这样抱皇上哎*

　　(She hugs the emperor like that!)

The source verbal message "do that" is explicitated into 抱皇上 ("to hug the emperor") in the subtitle present in the shot at 1:18:03 because of the visual image at 1:17:59 (Mulan hugs the emperor). Co-extension relation is thus created when "do" is extended into "hug" with this visual support. This co-extension relation changes to co-referentiality in Semiotic system II when "do that" is specifically transferred into 抱皇上 ("to hug the emperor") in the subtitle. The incorporation of the visual information in previous shot in the subtitle functions to modify the previous turn, in which another character (i.e. the emperor) is depicted. The narrative flow is thus maintained. The literal translation version 她被允许那么做吗 or 她能那么做吗 ("Is she allowed to do that?") fails to create this narrative flow.

Instance 6: [Scene at 27:24-27:25]

ST: *Your <u>serpentine salvation</u> is at hand!*

TT: *你的贴身护卫龙来了*

　　(Your own dragon guard is coming!)

A visual image of a dragon is shown in the shot at 27:24, and ST is present at 27:25. This source information "serpentine salvation" and the image of a dragon constructs a co-classification relation, while the translation version 护卫龙 ("dragon guard") at 27:25 creates a co-referentiality relation with the image. The previous turn (the shot at 27:24) is further modified because of this translation and the narrative flow is thus maintained.

The analysis above reveals roles and functions of semiotic interplay demonstrated through the implicit encoding device of anaphora and finds that co-referentiality is the most dominant cohesive tie between TT and image in Semiotic system II, which facilitates the maintaining of the narrative flow by means of modifying preceding turns. The next-subsection concentrates on influenced linguistic parts of subtitles because of semiotic cohesion.

3.2.2. Influenced linguistic parts

Table 3 above shows that linguistic parts relating to *identification* (11 out of 24 instances) and activity (7 out of 24 instances) are the most heavily influenced ones. This is as same as the case in cataphora. This is because visually shown participants and activity are the most prominent visual parts on screen, which then become the most easily captured information in audiovisual texts. What is different from the cataphoric case is that more linguistic expressions relating to *attributes* (5 out of 24 instances) are influenced. The reason behind it is that modifiers in preceding filmic units are more accessible to and easily received by viewers than those in subsequent units.

Two Instances are provided below for a further illustration. Instance 7 shows the influenced part relating to *attributes*, while Instance 8 concerns *identification*.

Instance 7: [Scene at 32:28-32:29]
ST: *Most impressive.*
TT: *还真不是普通的逊*
    (A group of losers!)

The visual image in the shot at 32:29 (undisciplined soldiers fighting with one another in a military camp and causing a great mess) justifies the translation of "Most impressive" into 还真不是普通的逊 ("A group of losers") at 32:38. The part "most impressive" indicates the participants' qualities: they are losers. The change of cohesive ties, from co-classification to co-referentiality exerts impacts on the translation of attributes.

Instance 8: [Scene at 45:54-46:05]

ST: *Hello! It's <u>the army!</u>*

TT: *拜托，这是<u>军书</u>哎！*

    (Come on, it's a military letter!)

In the shot at 45:54, it is visually shown that a letter is held in hand by a little dragon. At 46:05, the verbal message "the army" is specified into 军书 (a military letter) because of the visual image at 45:54. Thus, the linguistic part relating to identification is affected because of this visual-verbal cohesion.

Some patterns and regularities are revealed from the above analysis. A detailed discussion is provided in the next section, going along with a conclusion.

4. Discussion and conclusion

This paper, first, observes that subtitle translation is not limited to conveying information in the filmic unit of shots, in which source verbal messages, relevant visual images and subtitles are present synchronically, but it also draws on messages in the filmic units that are beyond shots, where the correlated visual images are not present in the same filmic units with source verbal messages and subtitles, but rather in previous or subsequent filmic units. This indicates subtitling units not only consist of shots, but also filmic units beyond shots such as scenes, sequences and stages.

Second, this paper finds that intermodal cohesive relations go through transformations in the subtitling process and the roles of the visual elements to the holistic communicative pattern also experience changes, though "visual ELEMENTS cannot normally be replaced during the process of translation" (Pérez-González 2014: 214, upper cases in original). This finding is contradictory to the proposal that visual elements' "contribution to the overall communicative gestalt in the target language remains unchanged" (Pérez-González 2014: 214). Through examining the different cohesive relations between images and ST in Semiotic system I and those between images and TT in Semiotic system II, different roles of the visual mode in relation to the spoken and written verbal mode are revealed. It is found that the prevailing cohesive relations in Semiotic system I is either co-classification, meaning that part of ST might be reduced or deleted because of the visual counterparts, or co-extension when implicit ST

needs to resort to images for further explanation. However, in Semiotic system II, the most predominant cohesive relation is co-referentiality, indicating that visual and verbal information identify to each other by referring to devices of reference. Different from linguistic references, such as pronouns, definite article and demonstratives, visual mode and written verbal mode identify to each other via correlating themselves to the four ideational features (*identification, activity, circumstances and attributes)*. Furthermore, it is found that, expressions concerning identification and activity are the more heavily influenced linguistic parts in subtitles in both cataphoric and anaphoric encoding devices.

Third, the significance of maintaining the narrative flow in subtitled films has been endorsed by the finding in this paper. It is noted that the narrative flow is usually maintained by means of reiterating participants or their activity in previous filmic units or heralding participants or their activity in subsequent units. This finding, on the one hand, validates the proposal that the understanding of the entire narrative structure in films is substantially mediated by characters, i.e. human identity (Mital et al. 2011) and non-human identity; and on the other hand, it reveals that the implicit information deriving from logical deduction (shown by activity) also benefits the maintaining of the narrative flow in subtitled films. The narrative flow in a film is like the coherence in a piece of writing, which is the way in which ideas in a text are linked logically, promoting the overall "understandability' of a text. The correlation of participants and activity in preceding and subsequent filmic units benefits the logical link among these units, thus facilitating the meaning-making process and promoting the coherence.

Though textual cohesion in audiovisual texts has been discussed in some previous studies (e.g. Valdés and Luque 2008), how communicative elements can be bound to realize cohesion in the subtitling process is still not be touched upon. For example, it has been proposed that "[c]ohesion is known as the property by which clauses or the components of a text hold together ⋯ by binding its surface elements together" (Valdés and Luque 2008: 135) and that "a filmic text is characterized by the cohesion between the visual and aural elements of the narration" (Valdés and Luque 2008: 136). Concerning how to realize cohesion in subtitle translation, this paper finds that cohesion between visual mode and written verbal mode can be realized by employing visual mode to reiterate ST in the subtitling process. This semiotic cohesion further facilitates the construction of the narrative flow in subtitled films.

Meanwhile, though cohesion of filmic texts has been discussed in some prior studies (e.g. Valdés and Luque 2008; Zabalbeascoa 1993), compositional meanings transferred in this process have not been highlighted. This paper stresses that intermodal cohesion in filmic units beyond shots is of great significance to indicating the holistic nature of audiovisual texts. It is suggested that a subtitled film is not a pasted collection of

fractions (i.e. shots), isolated from one another, but a holistic entity embodying semiotic interrelations among different filmic units. In this vein, the subtitling units are expanded from the conventionally recognized shots to filmic units beyond shots. Subtitled films become holistic entities.

Fourth, this paper further notes that though both of the two implicit encoding devices (anaphora and cataphora) function all the way in the subtitling process, the former is more actively involved than the latter. In previous studies (e.g. McNamara and Magliano 2009), the pointing forward reference is not taken into account. The finding in this paper foregrounds the importance of the pointing forward reference in subtitle translation and notes that this type of reference also plays a rather active role in setting up narrative structure in subtitled films.

The above-discussed points reveal that to maintain the holistic entity of subtitled films, compositional meanings in subtitle translation can never be ignored in the subtitling process.

References

Barthes, R. (1984). *Image-music-text* (S. Heath, Select &Trans.). London: Fontana Paperbacks.

Caple, H. (2013). *Photojournalism: A Social Semiotic Approach*. Basingstoke/New York: Palgrave Macmillan.

Chen, Y. and Wang, W. (2016). Relating visual images to subtitle translation in Finding Nemo: A multi-semiotic interplay. *Translation & Interpreting: The International Journal of Translation and Interpreting Research* 8(1): 69-85.

Chuang, Y. (2006). Studying subtitle translation from a multi-modal approach. *Babel* 52(4): 372-383.

Halliday, M.A.K. and Hasan, R. (1976). *Cohesion in English* (English Language Series 9). London: Longman.

Halliday, M.A.K. and Hasan, R. (1985). *Language, Context and Text: Aspects of Language in A Social-semiotic Perspective*. Melbourne: Deakin University Press.

Iedema, R. (2001). Analysing film and television: A social semiotic account of hospital: An unhealthy business. In van Leeuwen, T. and C. Jewitt (eds.), *Handbook of Visual Analysis. London*/Thousand Oaks/New Delhi: Sage Publications, 183-204.

Kress, G. and van Leeuwen, T. (2006). *Reading Images: The Grammar of Visual Design* (2nd edn.). London/New York: Routledge.

Lee, M., Beverly, R. and David R.E. (2013). The impact of subtitles on comprehension of narrative film. *Media Psychology* 16(4): 412-440.

McNamara, D. S. and Magliano, J. (2009). Toward a comprehensive model of comprehension. *The Psychology of Learning and Motivation* 51: 297–384.

Mital, K., Smith, J., Hill, L. and Henderson, M. (2011). Clustering of gaze during dynamic scene viewing is predicted by motion. *Cognitive Computation* 3: 5-24.

Pérez-González, L. (2014). *Audiovisual Translation: Theories, Methods and Issues*. London/New York: Routledge.

Remael, A. (2003). Mainstream narrative film dialogue and subtitling. *The Translator* 9(2): 225-247.

Royce, T. D. (2002). Multimodality in the TESOL classroom: Exploring visual-verbal synergy. *TESOL Quarterly 36*(2): 191–205.

Taylor, C. (2003). Multimodal transcription in the analysis, translation and subtitling of Italian Films. In Gambier, Y. (ed.), *Screen Translation: Special Issue of The Translator Studies in Intercultural Communication*. Manchester: St Jerome Publishing, 191-206.

Taylor, C. (2004). Multimodal text analysis and subtitling. In Ventola, E., Charles, C. and M. Kaltenbacher (eds.), Perspectives *on Multimodality*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 153–172.

Valdés, C. and Luque, F. (2008). Coherence in translated television commercials. *European Journal of English Studies* 12(2): 133-148.

Zabalbeascoa, P. (1993). *Developing Translation Studies to Better Account for Audiovisual Texts and Other New Forms of Text Production*. Unpublished PhD dissertation, University of Lleida.

Zhang, J. (2012). The interaction between visual and written ethnography in subtitling. *Visual Anthropology* (25): 439-449.

Authors' email addresses

Yuping Chen: ypc@cau.edu.cn; yche1694@uni.sydney.edu.au

Wei Wang: wei.wang@sydney.edu.au

About the authors

Dr Yuping Chen is a lecturer at China Agricultural University and her PhD is from the University of Sydney. Her research interests include translation studies, discourse analysis and second language acquisition.

Dr Wei Wang is a Senior Lecturer in School of Languages and Cultures at the University of Sydney, Australia. His research interests include discourse studies, sociolinguistics, translation studies and language education. His recent research concentrates on sociolinguistics and (critical) discourse analysis, especially on interdisciplinary studies of discourses in relation to contemporary China. His publications include Contemporary Chinese Discourse and Social Practice in China (Benjamins 2015) and Genre across Languages and Cultures (VDM Verlag 2007). His journal articles appear in Discourse Studies, Applied Linguistics Review, Journal of Multicultural Discourses, Journal of Chinese Language and Discourse, Australian Review of Applied Linguistics, Translation and Interpreting and other international academic journals. He also published book chapters with Mouton, Bloomsbury, Routledge, Benjamins, the University of Michigan Press, and Wiley-Blackwell.