Semiotic analysis of viewers' reception of Chinese subtitles: A relevance theory perspective

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

ABSTRACT

Against the backdrop of increased attention to semiotic relations in audiovisual translation, this paper examines how visual-verbal interplay influences subtitle translation and facilitates viewers' reception of subtitles. Drawing on semiotic analysis theories (e.g. Jewitt 2009; Thibault 2000; Taylor 2003) and relevance theory in Translation Studies (e.g. Sperber and Wilson 1995), this paper explores the visual-verbal interplay in a Chinese subtitled English language film *Night at the Museum: Battle of the Smithsonian*. The analysis reveals that when visual elements provide additional explanations to source verbal information, visual messages tend to be incorporated into subtitles. It suggests that this practice enhances the degree of relevance through aiding viewers' processing effort and boosting contextual effects. It also indicates that the integration of semiotic messages in subtitles contributes greatly to semiotic cohesion, which is instrumental in the construction of narrative flow in subtitled films. These insights and implications can enlighten both professional translators/subtitlers and academic researchers in the field of subtitle translation.

KEYWORDS

Semiotic interplay, relevance theory, contextual effects, processing effort, narrative flow.

1. Introduction

Semiotic analysis of subtitle translation has been of interest to many researchers (e.g. Chuang 2006; Kruger 2001; Oittinen 2008; Perego 2009; Taylor 2004; Tortoriello 2011). While roles of multimodal interplay in the subtitling process are analysed, previous studies mainly focus on semiotic relations between subtitles and visual images. Little effort has been made to address how visual-verbal interplay exerts impacts on viewers' comprehension of subtitles. Meanwhile, though some studies (e.g. Bogucki 2004; Tuominen 2011) have addressed viewers' comprehension of subtitles, analysis predominantly focuses on: (1) viewers' subjective accounts of the viewing strategies that were adopted in watching subtitled films, or (2) how visual-verbal relations work to ensure viewers' comprehension of film. However, the findings in the first type of studies might be influenced by many uncontrolled factors such as biased information or peer pressure, while in the second type, less attention has been given to other roles that

visual-verbal relations might play in facilitating viewers' interpretation of subtitles.

This paper aims to address viewers' reception of subtitled films from a relevance theory perspective through an exemplary case study. It analyses how semiotic interplay between visual and verbal modes affects the degree of relevance in subtitled films which, as argued, influences viewers' comprehension of subtitles. To this end, this paper presents a semiotic analysis of an American film with Chinese subtitles *Night at the Museum: Battle of the Smithsonian,* which is a popular English language fantasy-comedy film released in China.

This paper draws on semiotic analysis (multimodality) (Jewitt 2009; Thibault 2000; Taylor 2003) and relevance theory (Sperber and Wilson 1995) to analyse how visual-verbal relations work to improve the degree of relevance through maximising contextual effects and minimising viewers' processing effort of subtitles. It begins with an explanation of the data selection. Then, it moves on to the introduction of multimodality and relevance theory with the aim of construction of an analytical model for this study. Featuring a detailed data analysis on the subtitle translation of the sample film, this paper concludes with discussion of the analytical findings in relation to previous studies and provides suggestions and recommendations to professional translators/subtitlers and academic researchers.

2. Data: *Night at the Museum: Battle of the Smithsonian*

Night at the Museum: Battle of the Smithsonian was produced by Twentieth Century Fox Film Corporation and released in mainland China on 26 May 2009, with its box office reaching RMB119,000,000 (Movie box office ranking list in mainland China 2014). Interestingly, it is both an animated film and a live-action film. In live-action film, events are performed by actors or animals, which distinguishes it from actions performed by animated figures (Kuhn and Westwell 2012: 249). *Night at the Museum: Battle of the Smithsonian* has both fictional figures and figures represented by cast members. This hybrid film genre with multiple semiotic elements provides an excellent example for audiovisual translation research.

The simplified Chinese subtitle version of this film, authorised by the film production company (i.e. Twentieth Century Fox Film), is analysed in this paper. These Chinese subtitles are of great quality and have been widely accepted by Chinese film viewers. The subtitle translation in this film demonstrates how semiotic interplay increases the degree of relevance in the multimodal channel of communication and helps viewers to have a clear

and complete idea of subtitled films. In total, this paper identified 74 subtitle blocks in this film involving semiotic interplay for detailed semiotic analysis.

3. Theoretical foundations and framework

This section introduces two theoretical foundations, namely multimodality with a visual-verbal focus and relevance theory with a semiotic focus, with a view to constructing the theoretical framework for this study.

3.1 Multimodality: with a visual-verbal focus

Audiovisual texts are of great semiotic complexity in which different sign systems or modes, verbal and non-verbal, co-operate to create a coherent story. Subtitles are part of this multimodal system, interacting with and relying on all the film's different modes. This is because

> [t]he basic assumption that runs through multimodality is that meanings are made, distributed, received, interpreted and remade in interpretation through many representational and communicative modes – not just through language – whether as speech or as writing (Jewitt 2009: 14).

In this view, language is seen as part of a multimodal ensemble, but not regarded as the starting point of communication or the provider of a prototypical model of communication.

Drawing on Thibault's (2000) multimodal transcription, Taylor (2003) takes precedence to employ multimodal analysis in subtitle translation and argues that "the meaning potential of a film far transcends the spoken dialogue, and that any translation of film material should pay heed to the other semiotic modalities interacting with the verbal" (Taylor 2003: 194). This paper foregrounds how visual and verbal modes interact with each other and influence subtitle translation with a particular focal point on the impacts that this semiotic interplay brings about upon film viewers' reception of subtitles.

3.2 Relevance theory: with a semiotic focus

Relevance theory is a cognitive-pragmatic approach to communication proposed by Sperber and Wilson (1995), investigating how information is processed in a context. Relevance is a matter of degree, a relation between an assumption, referring to any representation treated as the true description of the actual world (Sperber and Wilson 1995: 74), and a context. "An assumption is relevant in a context if and only if it has some contextual effect in that context" (Sperber and Wilson 1995: 122). This indicates that contextual effect is one crucial factor in assessing the degree

of relevance of an assumption. In addition to contextual effects, "[t]he processing effort involved in achieving contextual effects" (Sperber and Wilson 1995: 122) is another factor impinging on the degree of relevance. Figure 1 below is compiled based on relevance theory proposed by Sperber and Wilson (1995) to illustrate the relation between contextual effects and processing effort.

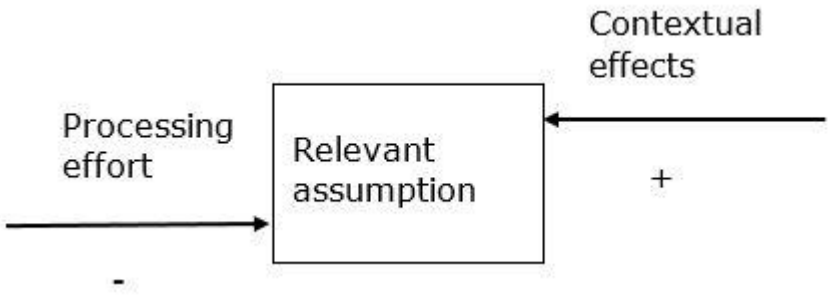

Figure 1. Contextual effects and processing effort

This figure shows the trade-off between contextual effects and processing effort. It means that "other things being equal, an assumption with greater contextual effects is more relevant; and, other things being equal, an assumption requiring a smaller processing effort is more relevant" (Sperber and Wilson 1995: 125). This is called optimal relevance, a principle playing a central role in ensuring a successful communication, in which utterance could enable addressees to locate the main meanings of the speakers without making unnecessary effort.

Relevance theory has long been used in the field of Translation Studies (e.g. Díaz-Pérez 2014; Jobes 2007), focusing on how verbal elements in contexts influence the degree of relevance. However, in the case of subtitle translation, visual and verbal modes work together to mediate contextual effects when viewers are exposed to visual images, source verbal dialogues and visually presented verbal subtitles simultaneously. In light of these facts, one interesting question arises. That is, how do verbal and non-verbal modes, especially the visual mode, interact with each other to enhance the degree of relevance through boosting contextual effects in subtitled films and alleviating viewers' processing burden?

3.2.1 Contextual effects

Prior to the discussion of contextual effects, it is necessary to pin down the definition and the main types of context in the field of Translation Studies. Context is defined by Sperber and Wilson (1986: 15) as "the set of premises used in interpreting it [an utterance]"; it is a "subset of the hearer's assumptions about the world". For Sperber and Wilson, context is a "cognitive environment", bound up with assumptions used by hearers to interpret utterances.

As to contextual effects, in line with the proposal that context is dynamic rather than static, and it is a dynamic process of interaction (Fetzer 2012), relevance theory holds that the underlying idea behind the notion of a contextual effect is that "[t]o modify and improve a context is to have effect on that context" (Sperber and Wilson 1995: 109), in which new information exerts impacts on old assumptions and gives rise to synthetic implication. Sperber and Wilson (1995: 114) hold that there are three types of contextual effects: "the addition of contextual implications," "the strengthening of previously held assumptions," and "the elimination of false assumptions" when "there is a contradiction between new and old information."

Given that subtitle is a mode of communication added to films, source verbal language and visual images are considered to be old assumptions whereas subtitles, as an addition to a finished film, are regarded as new information. How new information (i.e. subtitles) exerts impacts on old assumptions (i.e. source verbal information and visual images) has an essential bearing on contextual effects. Figure 2 below illustrates how source verbal messages, visual images and subtitles contribute to contextual effects.

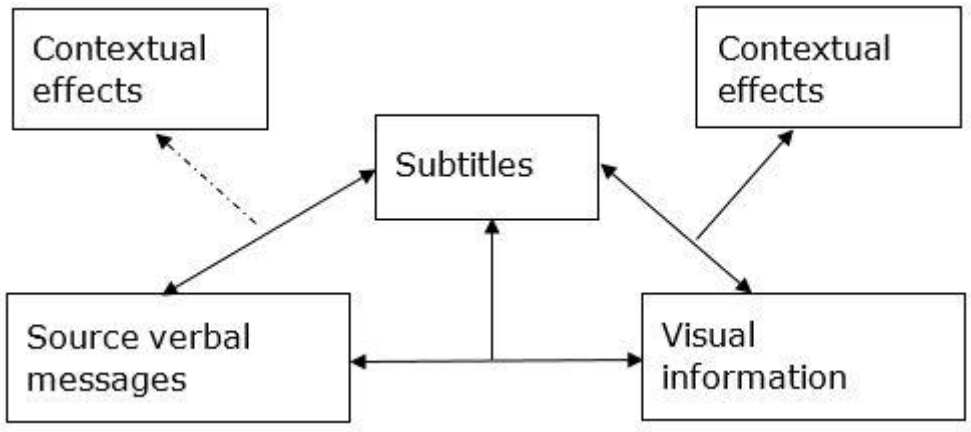

Figure 2. Formation of contextual effects in the subtitling process

Figure 2 shows that source verbal language does not directly contribute to the formation of contextual effects (see the dotted line). This is because most viewers of subtitled films are not able to catch the messages transmitted through the source language soundtrack. However, since source verbal language closely relates to subtitle translation, it is not advisable to skip source verbal language when discussing contextual effects. Also, the interplay between source verbal messages and visual information influences the translation of subtitles, and the intercommunication between visual information and subtitles directly contributes to contextual effects.

In order to analyse the interplay between source verbal messages and visual information, this paper identifies four types of visual-verbal relations.

The first is that of relay, where image and text "stand in a complementary relation" (Barthes 1984: 41), in which the image extends/advances the text and vice versa. The second is that of anchorage, meaning the text "directs the reader through the signifieds of the image, causing him to avoid some and receive others" by "elucidating" signs (Barthes 1984: 40), with signifieds referring to "a certain 'culture' of the society receiving the message" (Barthes 1984: 17). In anchorage, linguistic text occupies a predominantly important position over visual text. The third visual-verbal relation is redundancy, referring to when words and image communicate more or less the same information (Marleau 1982: 274 as cited in Díaz Cintas and Remael 2007: 50). To get a full picture of the interplay between visual and verbal modes, we draw on the understanding of anchorage and propose a fourth type of visual-verbal relation, namely moorage (Chen 2019: viii), in which visual messages further define and explicitate verbal information. The notion of moorage exists as a counterpart to anchorage: the former highlights the defining and explicitating function of visual to verbal mode, while the latter foregrounds the elucidating function of verbal mode to visual mode.

With regard to how the interplay between visual images and subtitles affects contextual effects, "the addition of new information which merely duplicates old information" does not count as an improvement of contextual effects (Sperber and Wilson 1995: 109), and a contextual effect is improved only when it involves the addition of new information which is related to old assumptions (i.e. visual images), the strengthening of old assumptions, or the elimination of contradictory previously held assumptions (Sperber and Wilson 1995: 109).

In addition to contextual effects, another key factor governing the degree of relevance is processing effort. Strong contextual effects do not inevitably lead to a high degree of relevance because more processing effort might be required. Strong contextual effects result in a high degree of relevance if and only if less processing effort is needed (Sperber and Wilson 1995: 124).

3.2.2 Processing effort

Viewers' processing effort of subtitles has been discussed extensively in previous studies (e.g. Koolstra *et al*. 2002; Lee *et al*. 2013). It has been argued that watching a subtitled film is more cognitively demanding and less effective in terms of content understanding and memory performance than viewing the same film in a dubbed version (Koolstra *et al*. 2002). One of the reasons is: "[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). Thus, how to

minimise the frequency of switching between subtitles and visual images is a key concern to alleviate viewers' processing effort.

Furthermore, some research has aptly noted that subtitle reading is largely automatic, requiring little additional cognitive effort (d'Ydewalle and De Bruycker 2007; d'Ydewalle and Gielen 1992) and more studies (e.g. Perego *et al*. 2010) also have found that a significant number of viewers' fixations and fixation times are devoted to the subtitled area rather than to the pictorial areas. In light of this, making subtitles represent and integrate as many messages as possible becomes a justifiable way to minimise the frequency of switching between subtitles and visual images. These messages include both the verbal information in source verbal dialogues and salient visual messages.

To that end, this paper differentiates two subtitling strategies correlated with viewers' processing effort. One is direct addressing (Chen 2019: 62), in which pictorial messages are incorporated into subtitles with source verbal information. The integration of visual and verbal information "may aid the reader insofar as he or she will not have to employ mental search strategies to retrieve (…) information" (Moran 2009: 55) from both visual information and subtitles. In this vein, less processing effort is required. The other is indirect addressing (Chen 2019: 62), in which subtitles and visual images transfer disconnected information: subtitles are solely dedicated to transmission of the source acoustic information, and the visual mode plays its role to deliver the visual information. In this case, viewers' eyes have to travel between the pictorial areas to the subtitled area to get a full understanding of intended messages. Relatively more processing effort is thus required.

3.3 Analytical framework for this paper

Drawing on the analysis of contextual effects and processing effort, the analytical framework used in this paper is constructed as shown in Figure 3 below.

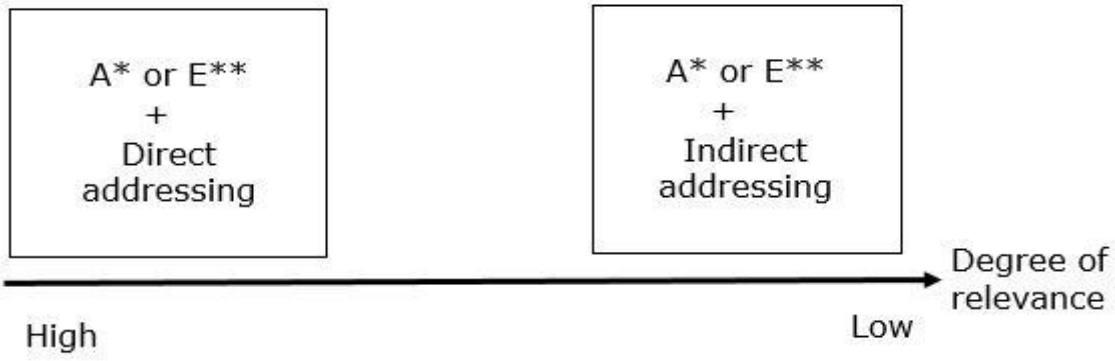

Figure 3. Degree of relevance in subtitle translation

Note*: A = Addition of new relevant information

Note**: E = Elimination of old contradictory information

Figure 3 shows that the degree of relevance is inextricably associated with two factors: contextual effects and processing effort. When new information relevant to old assumptions is added, information contradictory to old assumptions is eliminated or repetitive information to old assumptions is erased, strong contextual effects are created. Furthermore, if direct addressing is used in this process, requiring less processing effort, a high degree of relevance is set up. On the other hand, if indirect addressing is employed when new relevant information is added, old contradictory information or repetitive assumption is eliminated, relatively more processing effort is required and a low degree of relevance results.

4. Data analysis

The Chinese subtitles in the English language film *Night at the Museum: Battle of the Smithsonian* are examined by following the analytical framework above with the aim of discovering how semiotic interplay between source verbal information and pictorial messages influences subtitle translation and how subtitles interact with visual images to exert impacts on the degree of relevance and thus influence viewers' comprehension.

4.1 Contextual effects

In total, 74 instances involving semiotic interplay between source verbal language and visual images are identified and analysed, where 58 of them demonstrate a moorage relation, 14 reveal a redundancy relation, and two show a relay relation. In those 58 moorage instances, visual information provides relevant additional explanations to source verbal messages (see Table 1), while in those 14 redundancy instances, visual information justifies the elimination of the translation of part of the source verbal messages (see Table 2). In the two relay instances, visual and verbal information strengthen each other respectively (see Table 3).

1) *to do it* *做嘉宾* [to be the distinguished guest]	21) *contact* *点火* [lit it]	41) *Oh* *快看* [Look, a pretty.]
2) – *惹我发飙* [piss me off]	22) *it's hard* *你的头停不下* [your head cannot stop shaking]	42) *Again*. *又打* [Hit again.]
3) *shouldn't* *不该碰* [shouldn't touch]	23) *take* *你来驾驶* [You drive.]	43) *you* *你们* [you, plural form]
4) *it* *战斗* [battle]	24) *The doors*. *开门* [Open the doors.]	44) *you* *你们* [you, plural form]

5) *held down* *被关着* [being locked up]	25) *No*. *别跑* [Don't run.]	45) *boy* *没戏* [No use]
6) *it* *战争* [war]	26) *going* *躲着* [dodging]	46) *charge* *冲啊* [rush forward]
7) *finally* *总算接了* [answer it finally]	27) *these stars* *军衔* [military rank]	47) *this* *制服* [uniform]
8) *that* *拖我出来* [drag me out]	28) – *大脚* [big feet]	48) *There you go*. *下去吧* [Get down to the ground.]
9) *men* *兵* [soldier]	29) *second base* *打啵亲热* [kiss]	49) *we* *黑人* [black men]
10) – *山寨* [fake]	30) *you* *你们* [you, plural form]	50) *Torqueing angles* *小型机* [blimp]
11) *Finally* *总算到手了* [Got it finally.]	31) *it* *胸牌* [name tag]	51) *it* *摇摆* [shake]
12) *body* *胳膊……腿* [arms…legs]	32) *the Commons* *办公区* [office areas]	52) *stick* *操纵杆* [joystick]
13) *gun show* *二头肌* [biceps]	33) *bingo* *完事儿* [It's done.]	53) *Right there*. *站在那儿* [Stand there.]
14) *you* *你们* [you, plural form]	34) *you* *你们* [you, plural form]	54) *Streltsy* *士兵们* [soldiers]
15) *you* *你们* [you, plural form]	35) *bars* *信号* [signal]	55) *failure* *败将* [a defeated general]
16) *enough* *别吵了* [stop arguing]	36) *get* *借* [borrow]	56) *do* *扇* [fan]
17) *Streltsy* *士兵们* [soldiers]	37) *look* *听着* [listen]	57) – *交出* [hand over]
18) *in* *躲* [hide]	38) *get down* *蹲下* [squat]	58) *goes* *飞* [fly]
19) *get up* *站起来* [stand up]	39) – *笼子* [cage]	
20) *hold* *坐下不要动* [sit still]	40) *things* *沙漏* [sandglass]	

Table 1. Chinese subtitles based on moorage relation between source verbal messages and visual images

Table 1 lists all 58 instances in which visual information modifies or explicates source verbal information through providing additional explanations. Explicitation means that "[i]mplicit information in the ST [Source Text] is rendered explicit in the TT [Target Text]" (Munday 2016: 92). Implicit information in the original text is explicated in the translation information by resorting to the grammatical, semantic, pragmatic or discursive elements. Explicitation has been understood as one of the universal translation phenomena by many scholars (e.g. Chesterman 2004; Øverås 1998; Pápai 2004). In the domain of subtitle translation, because of the co-habitation of the verbal and the non-verbal, explicitation also takes

place when the specific meaning in the TT is derived from the non-verbal context in the ST. In these 58 Chinese translations, 29 are verbs and phrasal verbs, while another 29 are nouns and pronouns. Verbs and phrasal verbs are explicated by resorting to the relevant visual context to specify the action, for example *to do it* is rendered into *做嘉宾* [to be the distinguished guest] (i.e. 1), *shouldn't* into *不该碰* [shouldn't touch] (i.e. 3). Nouns or pronouns become more concrete because of visual elements, such as *this* is translated into *制服* [uniform] (i.e. 47), *it* into *胸牌* [name tag] (i.e. 31) and *things* into *沙漏* [sandglass] (i.e. 40). Furthermore, four Chinese translations solely depend on the visual information, i.e. *山寨* [fake] (i.e. 10), *大脚* [big feet] (i.e. 28), *笼子* [cage] (i.e. 39) and *交出* [hand over] (i.e. 57). All these instances demonstrate how the visual element functions to further define and clarify the verbal counterpart.

Salient visual participants are often referred to in subtitle translation. Being salient means "the materials which stand out to attract attention and direct viewers along certain paths of narrative construction" (Tseng 2013: 48). The salient position of participants can be acquired either "*immediately*" or "*gradually*" (Tseng 2013: 48, italicised in original). Film participants with immediate salience are presented to the audience in the foreground. Film participants' presentation can also be non-salient in the beginning and "be gradually 'upgraded' to a salient one" (Tseng 2013: 48). Based on the data analysis in this paper, it is further found that there is a third way to present salient participants, i.e. downgrading from an immediate salient one to a gradual salient one.

Contextual effects are strengthened when salient visual information is incorporated into subtitles to explicitate generic source verbal messages, culturally or semantically. First, drawing on the concept of homophora (Martin 1992: 126), identity retrieval from the context of culture, the non-verbal context of culture facilitates the retrieval of participants' identity, which, to a great extent, might be new and alien information to viewers of subtitled films. Thus, for viewers of subtitled films, contextual effects are strengthened due to the added new information relevant to old assumptions (see Example 1). Another case in which contextual effects can also be strengthened is when verbal context is semantically generic, but salient visual participants in the non-verbal context can justify its specific expression. Relevant new information is, thus, added and strengthened contextual effects are ensured (see Example 2).

ST: *It's over. It's over. It's all over.*
TT: *战争结束了!战争结束了!*
[The war is over! The war is over!]

Example 1. Transcription of the part of a scene[1] in *Night at the Museum: Battle of the Smithsonian* (38:04-38:21)

The participants (i.e. a marine and a nurse kissing each other [i.e. the victory kiss]) acquire their salience immediately upon presentation in the first shot[2]. Then, in the next few shots, this immediate salience downgrades to gradual salience: for example, either only part of the visual images (i.e. legs) are visually shown, only a background position is taken. These two (immediate to gradual) salient participants are also cross-modally presented in the visual text and in the generic source verbal text *it*. Drawing on the notion of homophora, it is much easier for American viewers to retrieve the identity of these participants from the American context of culture than Chinese viewers can and it is, thus, much easier for American viewers than Chinese viewers to infer that this scene relates to the victory of World War I. To bridge this cultural gap and ensure the intended comprehension of the original meaning, *it* is defined as the Chinese noun *战争* [war] in the Chinese subtitle. This rendering of information fills the cultural gap for Chinese viewers who might not be able to directly relate the victory kiss to the end of the war. This specified information does not entail any unnecessary meaning repetition, but reveals the gist of the meaning. Simultaneously, the Chinese noun *战争* [war] in the subtitle (i.e. new information) provides information interrelating with the visual image (i.e. old assumptions). In this way, a strong contextual effect is generated.

ST: *From the looks of things, I'd say he has a little over an hour.*
TT: *看着沙漏的情况他有一个多小时*
[From the looks of the sandglass, he has a little over an hour.]

Example 2. Transcription of the part of a scene in *Night at the Museum: Battle of the Smithsonian* (51:33-51:39)

What is visually presented is a man holding a sandglass saying, "From the looks of things, I'd say he has a little over an hour", while looking at the sandglass. The semantically generic source verbal expression *things* is specified as the Chinese noun*沙漏* [sandglass] in the subtitle. The rationale behind this translation is the presentation of the gradual salient visual participant (i.e. a sandglass) in this scene. Relevant new information (i.e. the Chinese noun *沙漏* [sandglass] in the subtitle) is added to old assumptions (i.e. visual image of a sandglass) and a strong contextual context is thus guaranteed.

The two examples above shed light on the fact that subtitle translation has much to do with the moorage relation between source soundtrack and visual images, which in turn boosts a strong contextual effect based on the interaction between visual images and subtitles.

Redundancy is another type of semiotic relation between visual images and source verbal messages, in which participants are identified both visually and verbally. Either the visual or the verbal is a definite and specific representation of the participant. Please see Table 2 below for the 14 instances showing this relation in *Night at the Museum: Battle of the Smithsonian.*

1) ladies and gentlemen	–	8) Mr. Daley	–
2) Mr. Daley	–	9) Mr. Daley	–
3) Teddy	–	10) fish	–
4) Lawrence	–	11) Mr. Daley	–
5) in front of you	–	12) Mr. Daley	–
		13) Mr. Daley	–
6) the hall	–	14) Dexter	–
7) Mr. Daley	–		

Table 2. Chinese subtitles based on redundancy relation between source verbal messages and visual images

In these instances, pictorial elements and source verbal language express more or less the same information. Visual images justify the elimination of the translation of the redundant source verbal information in subtitles. They are mainly interpersonal elements, such as greetings, interjections, vocatives, formulas of courtesy, etc. It is common practice to omit the translation of such elements in subtitling (Díaz Cintas and Remael 2007: 165). Contextual effect is strengthened because of the elimination of repetitive information (see Example 3).

ST: *No, you look, Mr. Daley.*
TT: *不 你听着*
[No, you listen.]

Example 3. Transcription of a shot in *Night at the Museum: Battle of the Smithsonian* (40:28)

A woman is addressing a man named Mr. Daley. The translation of the vocative *Mr. Daley* in the source verbal language is omitted in the subtitle, which does not entail any comprehension barrier to viewers because of the visual image of Mr. Daley on screen. The visual content, thus, fills the

linguistic gap left by the omission. The total information load does not suffer any losses. In contrast, the elimination of repetitive information contributes to a strong contextual effect.

The third type of semiotic interplay is relay. There are two instances in this respect in *Night at the Museum: Battle of the Smithsonian* (see Table 3).

1) slippers	*鞋* [shoe]
2) a rust hole in the crate	*板条箱的洞眼* [a hole in the wooden crate]

Table 3. Chinese subtitles based on relay relation between source verbal messages and visual images

Relay relation comes into being when source verbal information and pictorial information elucidate each other, meaning that neither of them can be self-sufficient to precisely deliver the intended meanings. In terms of contextual effects, contradictory or non-related information is added to old assumptions, so the contextual effect is weakened. One example is examined below (see Example 4).

ST: *Caught this one trying to escape through a rust hole in the crate.*
TT: *这小子想从板条箱的洞眼里逃走被抓了*
[Caught this one trying to escape through a hole in the wooden crate.]

Example 4. Transcription of the part of a scene in *Night at the Museum: Battle of the Smithsonian* (47:50-48:40)

A rust role in a container is saliently shown on screen, which does not precisely match the source verbal message *a rust hole in the crate*. These contradictories, or at least not precisely matching premises, entail a fair confusion in the translation: *crate* is translated into the Chinese noun *板条箱* [a wooden crate] and *a rust hole* is condensed to the Chinese noun *洞眼* [a hole]. The information transmitted in the subtitle matches neither the visual information nor the source verbal messages, meaning the new information neither adds relevant information, nor eliminates contradictory or repetitive information. In the wake, the strength of the contextual effect is weakened.

Based on the analysis of contextual effects, it is found that a moorage relation and a redundant relation between source verbal information and visual messages have the potential to improve contextual effects given that this interrelation provides subtitles the chance to add relevant information or eliminate repetitive information to visual images. This means subtitles

based on moorage and redundant cross-modal relations can provide a comparatively strong contextual context.

4.2 Processing effort

Many studies have explored how to facilitate viewers' processing of subtitles (e.g. Koolstra *et al.* 2002; Lee *et al*. 2013; Zhang 2012). Some of the studies conclude that switching between visual and verbal modes to collect information in subtitled films might add to viewers' processing burden. This paper proposes that narrowing down the number of communication modes exposed to viewers by direct addressing relieves viewers' burden of processing multimodal modes. Thus, less processing effort is required. On the other hand, when visual and verbal information is disconnected and transmitted respectively in their own communication channels, much processing effort is needed.

Number of instances	Subtitling	Processing effort required
58	Direct addressing	Less
16	Indirect addressing	More

Table 4. Subtitling strategies and processing effort

Table 4 shows that less processing effort is demanded in 58 out of 74 instances because of direct addressing (e.g. Example 5), while more processing effort is required in the remaining 16 instances due to indirect addressing (e.g. Example 6). It is interesting to note that the instances adopting the subtitling strategy of direct addressing are those based on a moorage relation between source verbal messages and images (i.e. as shown in Table 1 above), and the instances using the strategy of indirect addressing are those deriving from either a redundancy relation (i.e. as shown in Table 2 above) or those displaying a relay relation (i.e. see Table 3 above) between source verbal language and images.

ST: *I hate to ask, but, as you see, I'm missing a few body parts.*
TT: *我也不想麻烦你 可我缺胳膊少腿的*
[I hate to bother you, but I'm missing arms and legs.]

Example 5. Transcription of a shot in *Night at the Museum: Battle of the Smithsonian* (53:38)

The salient image of a bronze is presented on screen and speaks to two people, "I hate to ask, but…" Direct addressing is adopted in the subtitling process by integrating the gradual salient visual participant (i.e. a bronze [i.e. without arms and legs]) and the source verbal information (i.e. *missing a few body parts*) to generate the Chinese phrase *缺胳膊少腿的* [missing arms and legs] in the subtitle. Viewers' attention thus doesn't need to travel between the subtitled area and the pictorial areas. Since it has been proved that reading subtitles, to a great extent, is automatic, less processing effort is required when there is no need to travel to the pictorial areas. Meanwhile, the explicitation of the source verbal information also imposes new information on old assumptions, strengthening the contextual effects accordingly. So, a high degree of relevance is called for.

ST: *At the end of the hall, turn right…*
TT: *一直走到底，右拐*
[At the end of it, turn right.]

Example 6. Transcription of the part of a scene in *Night at the Museum: Battle of the Smithsonian* (23:54-23:56)

A man is walking in a hall and his son, consulting an online map, phones him and directs him to turn right at the end of the hall. The visual image of the hall justifies the elimination of the translation of the cross-modally repetitive source verbal message *the hall* in the subtitle. Though the contextual effect is strengthened in this way, viewers' eyes have to travel between the visual image and the subtitle for a complete understanding of the transmitted message. More processing effort is required. Therefore, a high degree of relevance cannot be guaranteed.

The above analysis verifies that direct addressing alleviates viewers' processing burden via narrowing down the number of communication channels that viewers have to refer to for information collection. This is made real through integrating source verbal messages with visual information, in which visual images are used to further interpret source verbal expressions and the explicitated information is delivered in subtitles. Thus, it can be noted that a moorage relation not only improves contextual contexts, but also requires less processing effort from viewers, which leads to a high degree of relevance. However, though a redundancy relation might improve contextual contexts through eliminating repetitive information to old assumptions, the entailed indirect addressing requires more processing effort. So, high contextual effects cannot be obtained.

5. Discussion and conclusion

This paper unravels that a moorage relation contributes to high contextual effects and that direct addressing requires less processing effort from viewers. So the integration of a moorage relation and direct addressing leads to a high degree of relevance in audiovisual texts. The rationale behind this argument is that semiotic cohesion is maintained in this process. As argued by Chaume (2004), cohesion in an audiovisual text operates on a semiotic, rather than on a merely semantic or lexicogrammatical level. Then how does semiotic interplay lead to semiotic cohesion? It "is very often achieved through reiteration between the verbal and the non-verbal" (Tortoriello 2011: 62) or synchrony between visual and verbal modes (Georgakopoulou 2009: 25; Di Giovanni 2003: 210). Moorage, in which visual messages further define and explicitate verbal information, realises the semiotic reiteration or synchrony, while direct addressing, in which pictorial messages are incorporated into subtitles with source verbal information, furthers this semiotic reiteration or synchrony. Thus, semiotic cohesion is achieved through the integration of a moorage relation and direct addressing in subtitling.

In another aspect, semiotic cohesion in turn contributes to maintaining the narrative flow in subtitled films. By its very nature, film dialogue "is not just 'dialogue', it is also a narrative" (Remael 2003: 233). As the translation of film dialogues, subtitles should have their own "sequential structure" (Remael 2003: 225) to rebuild the narrative flow for film viewers who cannot, or cannot fully, understand the source verbal language. In this sense, subtitle translation is, to a certain extent, a process of re-narrativising original films (Kruger 2010: 234). Since the invention of the concept of re-narrativisation (e.g. Kruger 2010: 234), its usage has been confined to audio narration for film viewers who are blind or partially sighted. In Kruger's study, it is argued that "the absence of codes from one of these semiotic systems means that the original narrative no longer operates in the same way and has to be re-narrativised in order for the audience to get the benefit of a coherent narrative" (2010: 231-232). In this vein, the re-narrativising process fills in the narrative gap to a sight impaired audience and complements a coherent story.

This paper further reveals that re-narrativisation also exists in subtitled films for ordinary viewers without sight problems. This is due to the fact that viewers of subtitled films do not have sufficient access to the messages transferred through the source verbal language, but need to rely on written subtitles as a bridge to help them comprehend subtitled films so as to maintain the narrative flow in subtitled films.

As to how the narrative flow in subtitled films can be maintained through re-narrativisation of the original films, this paper proposes that since subtitle belongs to a type of visual presentation — verbal sign (Delabastita 1989: 199), which is presented via a visual channel instead of an audio channel, the integration of visual message and source verbal information into the subtitle is a process of transforming the audio (+) visual narrative (Kruger 2010: 235) to a visual (+) visual narrative. This testifies to the justifiability of direct addressing, in which both visual information and source verbal information are verbally conveyed but visually presented through subtitles.

The above proposition on how the narrative flow in subtitled films can be re-narrativised and maintained through the adoption of moorage relation and direct addressing in subtitling sheds light on the following two implications for translators/subtitlers and researchers in the field of subtitle translation.

The professional implications that this insight brings about are twofold. Foremost, semiotic congruency does not always lead to omission, i.e. to delete the translation of the source verbal language which is congruent with the visual information. This goes against the commonly held existing proposition that visual images are usually expected to justify the deletion of part of or even the whole piece of source spoken information when visual images are congruent with source verbal language so as to overcome the temporal or spatial constraints in subtitling. This paper contends that incorporating congruent visual messages with the source verbal information in subtitles significantly contributes to viewers' comprehension of subtitled films as a consequence of narrowing down the number of multimodal communication channels exposed to viewers. This previous proposition and the finding in this paper have unravelled two major functions of congruent visual images in subtitling: to help overcome the technical constraints and to rebuild and maintain the narrative flow in subtitled films. As to which function is supposed to take a more pivotal position, there is no absolute answer. It all depends on which function is expected to be more active in different translation practices: to overcome the technical constraints is more significant or to maintain the narrative flow is more essential. Though it can never be said that either of them is unimportant, this paper contends that to sacrifice the narrative flow is at times the last resort if there is also the need to overcome technical constraints. This is because the readability and meaning transfer of subtitles will be completely lost if viewers do not have enough time to register the subtitles and there will be no way to guarantee viewers' comprehension of subtitles. But the proposition that the only function of visual images is to justify the deletion of the translation of source verbal information does not appear to be valid either.

Furthermore, another implication is that salient visual participants and visual activities exert more influential power than visual circumstances.

This paper noted that semiotic cohesion between subtitles and visual elements is mainly constructed in the following two ways: explicitating salient visual identification in subtitles and incorporating the information transmitted by key visual activity in subtitles. This means that the narrative flow in subtitled films is also mainly maintained through the above two ways, suggesting that translators/subtitlers should pay more attention to salient visual participants and activities as these two types of visual information are more likely to be involved in subtitle translation. Compared to identification and activity, circumstances, especially the visual circumstances in films, do not take an equally pivotal position in subtitling.

Another implication of the findings in this paper concerns how to take optimal advantage of visual information and its interactions with verbal counterparts which has become a crucial issue in this newly developing field of subtitle translation. Since subtitling is such a complex translation activity involving so many facets, it is logically expected to either respectively construct an analytical framework for every single issue or build up a comprehensive model intended to cover all the interweaving issues in subtitling. Though the latter might only be an aspiration at this stage, it is extremely helpful to try to correlate the findings in various studies with the analytical models for different single issues so as to propel the construction of a comprehensive model at last. It is hoped that this paper has provided insights for the construction of an analytical model to address viewer's reception of subtitles and that the findings in this study can be correlated with previous studies on overcoming the technical constraints in subtitle translation.

References

- **Barthes, Roland** (1984). *Image Music Text.* Selected and translated by Stephen Heath. London: Fontana Paperbacks.

- **Bogucki, Łukasz** (2004). *A Relevance Framework for Constraints on Cinema Subtitling*. Łódź: Wydawnictwo Uniwersytetu Łódzkiego.

- **Chaume, Frederic** (2004). *Cine Y Traducción* [Cinema and Translation]. Madrid: Cátedra.

- **Chen, Yuping** (2019). *Translating Film Subtitles into Chinese*. *A Multimodal Study.* Singapore: Springer.

- **Chesterman, Andrew** (2004). "Beyond the particular." Anna Mauranen and Pekka Kujamäki (eds) (2004). *Translation Universals. Do they exist?* Amsterdam/Philadelphia: John Benjamins, 33–49.

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

- **Delabastita, Dirk** (1989). "Translation and mass-communication: Film and TV translation as evidence of cultural dynamics." *Babel* 35(4): 193-218.

- **Díaz Cintas, Jorge and Aline Remael** (2007). *Audiovisual Translation: Subtitling*. Manchester: St. Jerome Publishing.

- **Díaz-Pérez, Francisco Javier** (2014). "Relevance theory and translation: Translating puns in Spanish film titles into English." *Journal of Pragmatics* 70: 108-129. doi: http://dx.doi.org/10.1016/j.pragma.2014.06.007.

- **Di Giovanni, Elena** (2003). "Cultural otherness and global communication in Walt Disney films at the turn of the century." Yves Gambier (ed.) (2003). *Screen Translation. Special Issue of The Translator. Studies in Intercultural Communication* 9(2): 207-223.

- **d'Ydewalle, Géry and Ingrid Gielen** (1992). "Attention allocation with overlapping sound, image, and text." Keith Rayner (ed.) (1992). *Eye Movements and Visual Cognition. Scene Perception and Reading*. New York: Springer, 415-427.

- **d'Ydewalle, Géry and Wim De Bruycker** (2007). "Eye movements of children and adults while reading television subtitles." *European Psychologist* 12: 196-205. doi: 10.1027/1016-9-40.12.3.196.

- **Fetzer, Anita** (2012). "Context in interaction." Rita Finkbeiner, Jörg Meibauer and Petra B. Schumacher (eds) (2012). *What is a Context? Linguistic approaches and challenges*. Amsterdam/Philadelphia: John Benjamins, 105–128.

- **Georgakopoulou, Panayota** (2009). "Subtitling for the DVD industry." Jorge Díaz Cintas and Gunilla Anderman (eds) (2009). *Audiovisual Translation: Language Transfer on Screen*. Palgrave Macmillan, 21-35. doi: 10.1057/9780230234581.

- **Iedema, Rick** (2001). "Analysing film and television. A social semiotic account of *Hospital: an Unhealthy Business*." Theo Van Theo and Carey Jewitt (eds) (2001). *The Handbook of Visual Analysis*. London: Sage, 183-204.

- **Jewitt, Carey** (2009). "An introduction to multimodality." Carey Jewitt (ed.) (2009). *The Routledge Handbook of Multimodal Analysis.* London/New York: Routledge, 14-27.

- **Jobes, Karen H.** (2007). "Relevance theory and the translation of scripture." *Journal of the Evangelical Theological Society* 50(4): 773-797.

- **Koolstra, Cees M., Allerd L. Peeters and Herman Spinhof** (2002). "The pros and cons of dubbing and subtitling." *European Journal of Communication* 17(3): 325–354. doi: 10.1177/0267323102017003694.

- **Kruger, Helena** (2001). "The creation of interlingual subtitles: Semiotics, equivalence and condensation." *Perspectives: Studies in Translatology* 9(3): 177-196. doi: 10.1080/0907676X.2001.9961416.

- **Kruger, Jan-Louis** (2010). "Audio narration: re-narrativising film." *Perspectives: Studies in Translatology* 18(3): 231-249. doi: 10.1080/0907676X.2010.485686.

- **Kuhn, Annette and Guy Westwell** (2012). *A Dictionary of Film Studies*. Oxford: Oxford University Press. doi: 10.1093/acref/9780199587261.001.000.

- **Lee, Mina, Beverly Roskos and David R. Ewoldsen** (2013). "The impact of subtitles on comprehension of narrative film." *Media Psychology* 16(4): 412-440. doi: 10.1080/15213269.2013.826119.

- **Marleau, Lucien** (1982). "Les sous-titres...un mal nécessaire" [Subtitles...a necessary evil]. *Meta*. *Translators' Journal* 27(3): 271-285.

- **Martin, James R.** (1992). *English Text: System and Structure*. Philadelphia/Amsterdam: John Benjamins.

- **Moran, Siobhan** (2009). *The Effect of Linguistic Variation on Subtitle Reception*. Master's thesis. York University. https://scholar.google.co.nz/scholar?hl=en&as_sdt=0%2C5&q=the+effect+of+linguistics+variation+subtitle+reception&btnG=#d=gs_qabs&u=%23p%3DFlp_fcJApuQJ (consulted date 12.02.2018).

- **Movie box office ranking list in mainland China (内地电影票房排行榜)** (2014). http://58921.com/alltime (consulted date 12.12. 2014).

- **Munday, Jeremy** (2016). *Introducing Translation Studies. Theories and Applications*. 4th ed. London/New York: Routledge.

- **Oittinen, Riitta** (2008). "From Thumbelina to Winnie-the-Pooh: Pictures, words, and sounds in translation." *Meta. Translators' Journal* 53(1): 76-89. doi: https://doi.org/10.7202/017975ar.

- **Øverås, Linn** (1998). "In search of the third code: An investigation of norms in literary translation." *Meta. Translators' Journal* 43(4): 571-588. doi: https://doi.org/10.7202/003775ar.

- **Pápai, Vilma** (2004). "Explicitation: A universal of translated text?" Anna Mauranen and Pekka Kujamäki (eds) (2004). *Translation Universals: Do they exist?* Amsterdam/ Philadelphia: John Benjamins, 143-164.

- **Perego, Elisa** (2009). "The codification of non-verbal information in subtitled texts." Jorge Díaz Cintas (ed.) (2009). *New Trends in Audiovisual Translation.* Bristol: Multilingual Matters, 58-69.

- **Perego, Elisa, Fabio Del Missier, Marco Porta and Mauro Mosconi** (2010). "The cognitive effectiveness of subtitle processing." *Media Psychology* 13: 243-272. doi: 10.1080/15213269.2010.502873.

- **Remael, Aline** (2003). "Mainstream narrative film dialogue and subtitling." *The Translator* 9(2): 225-247. doi: 10.1080/13556509.2003.10799155.

- **Sperber, Dan and Deirdre Wilson** (1986). *Relevance. Communication and Cognition.* Oxford/Cambridge: Blackwell.

- **Sperber, Dan and Deirdre Wilson** (1995). *Relevance. Communication and Cognition.* 2nd ed. Oxford/Cambridge: Blackwell.

- **Taylor, Christopher J.** (2003). "Multimodal transcription in the analysis, translation and subtitling of Italian films." Yves Gambier (ed.) (2003). *Screen Translation. Special Issue of The Translator. Studies in Intercultural Communication* 9(2): 191-206.

- **Taylor, Christopher J.** (2004). "Multimodal text analysis and subtitling." Eija Ventola, Cassily Charles and Martin Kaltenbacher (eds) (2004). *Perspectives on Multimodality*. Amsterdam/Philadelphia: John Benjamins, 153-172.

- **Thibault, Paul J.** (2000). "The multimodal transcription of a television advertisement: Theory and practice." Anthony Baldry (ed.) (2000). *Multimodality and Multimediality in the Distance Learning Age.* Campobasso: Palladino Editore, 311-384.

- **Tortoriello, Adriana** (2011). "Semiotic cohesion in subtitling: The case of explicitation." Adriana Serban, Anna Matamala and Mean Marc Lavaur (eds) (2011). *Audiovisual Translation in Close-up. Practical and Theoretical Approaches*. Bern: Peter Lang, 61-74.

- **Tseng, Chiaoi** (2013). *Cohesion in Film. Tracking Film Elements.* Hampshire: Palgrave Macmillan. doi: 10.1057/9781137290342.

- **Tuominen, Tiina** (2011). "Accidental reading? Some observations on the reception of subtitled films." Adriana Serban, Anna Matamala and Mean Marc Lavaur (eds) (2011). *Audiovisual Translation in Close-up. Practical and Theoretical Approaches*. Bern: Peter Lang, 189-204.

- **Zhang, Jinghong** (2012). "The interaction between visual and written ethnography in subtitling." *Visual Anthropology* 25: 439-449. doi: 10.1080/08949468.2012.720200.

Biographies

Yuping Chen is an Associate Professor 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. She is the author of *Translating Film Subtitles into Chinese: A Multimodal Study* (2019). Her publications have appeared in *Chinese Translators Journal*, *T&I Review*, *Translation & Interpreting: The International Journal of Translation and Interpreting Research* and other academic journals.

E-mail: yuping-chen@outlook.com / ypc@cau.edu.cn

Wei Wang is a Senior Lecturer in Translation Studies at the University of Sydney. His primary research interests are in the areas of Discourse Studies and Translation Studies. He is the author of *Media Representation of Migrant Workers: Identities and Stances* (2017) and *Genre across Languages and Cultures* (2007). His publications have appeared in *Discourse Studies, Applied Linguistics Review, T&I Review, Journal of Multicultural Discourse* and other international academic journals. He has also published book chapters with Routledge, Continuum, Benjamins, the University of Michigan Press, Mouton, and Wiley-Blackwell.

E-mail: wei.wang@sydney.edu.au

Notes

[1] Scenes "comprise more than one shot. The defining characteristic of scenes is their continuity of time and space" (Iedema 2001: 188). All shots in a same scene remain in one time-space. In this paper, if an example in a scene is referred to, it means we analyse more than one shot in one time-space.

[2] In a shot the camera movement is unedited (uncut): "[i]f the camera's position changes, this may be due to panning, tracking, zooming, and so on, but not editing cuts" (Iedema 2001: 189). Examples of a shot analysed in this paper concern a single still without cuts.