

Cross-Modal Translation: A Study of English Subtitle Translation from the Perspective of Multimodal Interaction

Rongbin Wang^{*}, Yaoqin Xue, Haizhen Wei

School of Foreign Language Shanxi University, Taiyuan, China

Email address:

wangrongbin@sxu.edu.cn (Rongbin Wang), xueyaoqin@sxu.edu.cn (Yaoqin Xue), 2576552369@qq.com (Haizhen Wei)

^{*}Corresponding author

To cite this article:

Rongbin Wang, Haizhen Wei, Yaoqin Xue. Cross-Modal Translation: A Study of English Subtitle Translation from the Perspective of Multimodal Interaction. *International Journal of Applied Linguistics and Translation*. Vol. 8, No. 1, 2022, pp. 1-9.

doi: 10.11648/j.ijalt.20220801.11

Received: December 7, 2021; **Accepted:** December 27, 2021; **Published:** January 8, 2022

Abstract: Under the guidance of the Systemic-functional synthetic framework for multimodal discourse analysis, this paper collects multimodal corpus from the classical American TV show *Friends* and undertakes a thorough qualitative analysis of how cross-modal relations influence the Chinese translation of English subtitles. As is shown, in order to give Chinese viewers a better understanding of English subtitles in multimodal contexts, different translational approaches should be used for the translation of subtitles that are in different cross-modal relations, which are categorized into two general types, that is, complementary relation on the one hand and non-complementary relation on the other hand. The qualitative analysis in this paper finds that though strategies of translation that are appropriate to make use of in the process of English translation into Chinese include amplification, paraphrasing, annotation, direct translation, free translation, and linear translation, the choice of translation approach is directly influenced by cross-modal relations. For example, for subtitles that belong to multimodal discourse of reinforcing complementary relation, it is advisable for translators to make use of translational strategies of amplification, paraphrasing and annotation to highlight the foregrounded meaning expressed by the dominant modality on the one hand, and on the other hand to explicitly and verbally express the background information provided by the backgrounded modality if discourses of two modalities in the same scene form a relation of highlighting. As cross-modal communication becomes continuously widespread and the popularity of English shows keep arising in China, this paper endeavors to provide some useful implications for the translation of English subtitles, which as verbal and auditory discourse is part of a larger multimodal discourse in films.

Keywords: Multimodal Interaction, Subtitles, Chinese Translation

1. Introduction

Multimodal Discourse Analysis (MDA), as a new turn of discourse analysis originally developed in the west [11], has been drawing attention from both linguistic and social semiotic fields. As is defined by Kress & van Leeuwen, two leading scholars in MDA, “multimodal discourse” is a kind of discourse in which multiple types of semiotic modals including sound, text and image etc. are integrated to convey meaning and express communicative purposes [2]. An analytical framework for visual discourse based on Halliday’s Systemic-Functional Linguistics was developed [1]. Some Chinese scholars tried to build their localized theories and

developed analytical approaches related to multimodal discourse analysis. For example, Zhang [14] proposed to undertake multimodal analysis from a comprehensive perspective that took the influential factors of culture, context, content, and expression into consideration and built a systemic-functional synthetic framework for multimodal discourse analysis. Li [4] also contributed to the development of multimodal discourse analysis within the guide of Systemic-Functional Linguistics [14].

Multimodal interactional analysis is a new and important branch of multimodal discourse analysis, which concerns about two analytical aspects: one emphasizes the multimodal nature of human interaction and the other undertakes research of human interaction through a comprehensive investigation

of various modals [7]. This current study belongs to the former. Relations among discourses of different modalities in a multimodal discourse can be divided into two categories: intersign and resymbolization [15]. In the study of intersign relations, Zhang [15] classified intersign relations into two types with one being complementary and the other being non-complementary, based on the role played by discourses of various modes in constructing overall meaning.

Film is a form of multimodal discourse that integrates various semiotic modalities such as sound, image, subtitle, and color, etc. The exploration of translational processes of film subtitles and of strategies involved is inseparable from multimodal discourse analysis of film discourse. Previous studies on film subtitle translation mainly adopted theoretical approaches such as Skopos theory, domestication and foreignization, Functional Equivalence theory, Relevance Theory as well as doing analysis from aesthetic and cultural perspectives [5]. As for multimodal translation studies, most previous studies focus on corpus like films, TV programs, and most literature adopted the theoretical framework of multimodal analysis proposed by Zhang [15] and analyzed from the four aspects of culture, context, content, and expression [12]. Such multidimensional analysis of original texts is helpful for translators in that a better understanding of original texts through an examination of the interactional relations between different modalities in the same text can assist translators to select translation strategies that are more appropriate for specific multimodal texts.

This paper first tries to examine the interactional relations among subtitles as texts, characters' facial expressions, intonations, auditory reactions of the audience on the spot, etc. in multimodal corpus that are excerpted from the classical American TV series *Friends* from the perspective of multimodal interaction, and then to explore and summarize corresponding Chinese translation strategies of English subtitles, hoping to provide some new insight into the study of multimodal English to Chinese subtitle translation.

2. Theoretical Framework

There are two kinds of interactional relations between different modalities: complementary relation on the one hand and non-complementary relation on the other hand [15]. In multimodal discourses where the use of mono-modality does not fully express its meaning and is supplemented by another modality, the relationship between these two modalities involved is defined as being complementary. To the opposite, non-complementary means that discourse of a modality, though still contributes to the process of meaning-making by discourse of another modality, only plays an insignificant role in cross-modal meaning-making [15]. Complementary relation is further classified into reinforcing complementary and non-reinforcing complementary [15]. Within reinforcing complementary, there are three specific kinds of relations: highlighting relation in which discourse of the foregrounded modality is provided background information by other discourse of other modalities that highlight the foregrounded

one, primary-secondary relation in which discourse of a modality plays an assisting role in the process of meaning making by discourse of another modality, and relation of expansion in which discourse of one modality extends the meaning expressed by discourse of another modality [15]. Within non-reinforcing complementary, there are also three specific kinds of relations: coordinating relation in which discourses of different modalities jointly express the overall meaning and the absence of discourse of any modalities leads to an incompleteness of the overall meaning of the multimodal discourse as a whole, united relation in which different types of media form a joint relation in the same modality, and intersecting relation in which discourses of two modals represent the overall meaning alternately [15]. Within non-complementary relation, there are again three types of relations: first, overlapping relation which further includes redundant relation in which one modality is only used for the sake of accessible reading, not strengthening nor supplementing the other modality, exclusionary relation in which there is a contradiction between one modality and others, counteracting relation in which one modality is irresistible to the other. In other words, two modalities counteract with each other; second, inclusive relation which is further classified into whole-part relation, abstract-concrete relation; and third, contextual-interactional relation which consists of independent relation and dependent relation [15].

In previous studies of subtitle translation, translational strategies frequently used include reduction, amplification, free translation, conversion, literal translation, domestication, and foreignization, etc. Reduction means the deletion of irrelevant information within the limited time and space of the information receiver [6]. Amplification refers to the need to add some amplification information to help viewers better understand the original text if literal translation cannot accurately convey the original information [13]. Free translation means not to translate word by word. Conversion refers to the change of word classes in translation that Chinese uses verbs more often while English prefers nominal expressions, which determines the conversion of parts of speech involved in translation, and the conversion between English abstract words and Chinese concrete words, etc. [9]. Literal translation is to translate the original content directly. The strategy of domestication and foreignization is proposed by Venuti [3], an American translator, while domestication requires translators to approach the target readers and integrate the original into the culture of the target language, the strategy of foreignization considers the differences between Chinese and Western cultures and appropriately retains the cultural characteristics of the source language [6].

3. Corpus Collection

This paper takes multimodal excerpts from season 1 of *Friends*, a classic American TV sitcom, as its multimodal corpus. *Friends* is a sitcom that centered on funny stories happened to six close friends who lived in Manhattan, New York, and tells humorous stories that happened in their

ten-year living together. The function of multimodal interaction is particularly obvious in this show. Discourses of various modalities such as visual and auditory-verbal modalities cooperate with one another and make the show vivid and authentic. By analyzing the relations between different modalities including relations between subtitles as auditory-verbal modality and performances, setting, etc. as visual modality, this paper tries to figure out some appropriate strategies of English subtitles translation into Chinese. To be specific, this paper collects representative multimodal corpus from twenty-four episodes of season 1 of the show and identifies the interaction between subtitles and discourse of other modalities occurring in the same scene. Among the thirteen kinds of modal relations [14], this paper finds no cases of abstract-concrete relation and dependent-independent relation. In other words, this paper undertakes specific analysis of other eleven kinds of modal relations.

4. Corpus Analysis

4.1. Translation Strategies Regarding Multimodal Discourse of Complementary Relation

4.1.1. Translation Strategies Regarding Multimodal Discourse of Reinforcing Complementary



Figure 1. Example 1 (Episode 1. 07:23).

In this scene, the woman facing the camera in the middle of the picture is named Rachel, one of the leading characters of the show, who escaped from her own wedding to her friend Monica's place. In this case, the two characters' conversation, that is, the auditory-verbal discourse, provides background information about what was happening and therefore was backgrounded rather than foregrounded. On the contrary, what is highlighted and foregrounded in this case is the character Rachel's facial expression. The character's sad facial expression as visual discourse was designed to be centralized in the scene, making the expression prominent and obvious to viewers. In other words, this nonverbal modality was foregrounded so that viewers could first tell that the character

was deeply sad even without referring to what the conversation was about or what the subtitles read. About the linguistic features of the two sentences from friend Monica, *Okay, look, this is possible for the best, you know?* and *Independence. Taking control of your life.*, there are both pronouns *this* that refers to contextual information about what happened earlier in the church, and zero subject sentence that also requires contextual interpretation. Both the foregrounded-backgrounded relations between nonverbal facial expressions and auditory-verbal subtitles, and the context-dependent characteristics of subtitles determines that verbal subtitles in this case are secondary compared to the centralized facial expression. In the process of translation to Chinese, it is advisable to make use of the strategy of omission and amplification to translate the first sentence:

Source: Okay, look, this is probably for the best, you know?

Translation: 好了, 别难过了, 或许逃婚反倒最好。

hǎo le, bié nán guò le, huò xǔ táo hūn fǎn dào zuì hǎo.

In detail, on one hand, the expressions of "Okay", "you know" in this scene do not play a role in highlighting the meaning conveyed in facial expression therefore, the modal particle and the discourse marker are translated as an individual phrase 好了 (*Ok*), and on the other hand, the meaning expressed in foregrounded facial expression, that is, Rachel being very sad, is also taken into consideration in translation, 别难过了 (*Don't be upset.*) was added into Chinese translation. By doing so, the information conveyed in non-verbal discourse is also expressed verbally in target language, making the scene easier to understand. Similarly, for the translation of the second sentence, the strategy of amplification is also adopted. To be specific, in the target language "Independence. Taking control of your life" is translated as a complete sentence, which makes the verbal reference more clearly and directly map to centralized character Rachel whose facial expression was foregrounded.

Source: Independence. Taking control of your life.

Translation: 你独立了, 可以主宰自己的人生。nǐ dú lì le, kě yǐ zhǔ zǎi zì jǐ de rén shēng.



Figure 2. Example 2 (Episode 4. 05:41).

Source: And while we're on the subject of news.

Translation: 人家现在也是好事将近哟。

rén jiā xiàn zài yě shì hǎo shì jiāng jìn yō.

In this scene in *Example 2*, the woman on the left of the picture was engaged and she was in a casual conversation with her friend. However, viewers may get confused if they only pay attention to what the friend in black coat said and what is read in the subtitle "And while we're on the subject of news.", because the auditory-verbal discourse does not provide

enough information about what the “news” is about. In this case, the engaged woman’s gesture as of showing the engagement ring on her finger, as a visual discourse on the one hand and as a representational gesture that entails semiotic meaning [10], plays a supplementary role to the auditory-verbal discourse. In other words, it helps viewers understand what the woman was talking about by giving additional visual contextual information. The visual discourse without doubt narrows down viewers’ explanation of the subtitle as auditory-verbal discourse. In translation, since visual discourse plays a secondary role to the auditory-verbal discourse, it is necessary to verbalize the contextual information provided by the visual discourse to clarify the meaning of the original sentence, the strategy of paraphrasing is adopted. In detail, translation here is not confined to the linguistic form of *on the subject* which means “on a certain topic”. Rather, it is translated as 好事将近 (Fortunate things are approaching.), which actually combines the information provided by the visual discourse (ring on the finger). Besides, *we* is translated as 人家 (*my friend*) here which in Chinese language not only refers to the first person plural, but also is an intonation, which highlights the bragging feeling of the woman on the left of the picture who was about to get married. In this case, again the consideration of multimodal interaction helps translators to build an appropriate Chinese translation that helps Chinese viewers understand the plot better.



Figure 3. Example 3 (Episode 18. 22:20; 22:24).

In this scene, visual discourse expands the meaning expressed by what the characters said as auditory-verbal discourse. In detail, the friends were playing the game of draw-and-guess in which the female character Rachel was drawing, and other friends were guessing the name of related movies. Within visual discourse, Rachel drew an image of a bean while saying *bean-bean* loudly. In this case, visual discourse is a supplement to auditory-verbal discourse, forming

a relation of expansion. Right after Rachel finished her drawing, one of the male friends Joey shouted out the correct answer, that is, *The Unbearable Lightness of Being*. However, the logical and cultural connection between beans and the movie’s title was far easy to be understood by Chinese viewers. If only a translation of the mentioned movie’s title is provided, viewers will still not understand the humorous meaning of this scene. Considering this, the strategy of annotation is adopted in the process of translation to give viewers an additional information about the relationship between *beans* as are drawn on the whiteboard and spoken out by the female character and *being* as is in the title of the movie, that is, the words *bean* and *being* share similar pronunciation. This annotation is suggested to be further added to the top of the scene.

Source: Bean! Bean! The unbearable lightness of being!

Translation: 豆子! 豆子! 《布拉格之恋》! (注释: 电影名中的 being 和 bean 发音相似)

dòu zǐ! dòu zǐ! bù lā gé zhī liàn! (Annotation: bean and being share similar pronunciation.)

As is shown above, while translating English subtitles into Chinese, it is advisable to adopt the strategies of amplification, paraphrasing, and annotation both to highlight the foregrounded meaning expressed by the dominant modality, and to explicitly express the background information provided by the backgrounded modality if discourses of two modalities in the same scene form a highlighting relation. However, it is appropriate to make use of the strategy of free translation or paraphrasing to present the meaning expressed in the discourse of secondary modality when discourse of two modalities in the same scene are in the relation of primary-secondary relation. As for cases when discourse of two modalities belong to relation of expansion, the strategy of annotation can be adopted in translation.

4.1.2. Translation Strategies Regarding Multimodal Discourse of Non-reinforcing Complementary



Figure 4. Example 4 (Episode 3. 15:21).

In this scene, there is a coordinating relation between visual and auditory discourses. In other words, visual and auditory discourses work together in expressing the overall meaning. In detail, this scene mainly expresses an ironical meaning, which would be incomplete without the use of discourse of either modality. Specifically, the context in which the scene took place roughly like this: the male character Chandler (the man on the left in the picture above) couldn't quit smoking, so his friends stared at him with disapproval and dislike, but Chandler retorted and argued that everyone had flaws and that there should be mutual acceptance and tolerance among friends. Later, the friends present began to mock one another's flaws. In this process, individual characters' rhetorical questions, counter-talking tones, and facial expressions expressing ironical meaning combine to convey the sense of humor. The speakers' verbal expression (auditory modality), their facial expressions while speaking (visual modality), and the laughter of present viewers in the background come to convey the humorous meaning by coordinating with each modality. The absence of any one of these modalities will reduce the humorous meaning conveyed accordingly. When the character Joey commented on the comical behavior of the character Phoebe biting her own hair, another character, Ross, thought the behavior was cute, so Joey asked back, *Oh, you do, do you?* In the translation process, instead of translating it literally, the strategy of paraphrasing is adopted and this line is translated as 哦? 你不会真的认为这样可爱吧? (*You don't really think Phoebe's behavior is cute, do you?*). Taking multimodal coordination into consideration, this translation not only preserves the speaker's rhetorical and mocking tone, but also retains the meaning of multimodal humor by paraphrasing. When the character Ross defended that her biting did not break the law, Rachel replied sarcastically *Indeed, there isn't*. In translating this line, the researcher does not translate it literally either. Instead, this line is also paraphrased and translated as 揪头发确实没犯法 (*Indeed you do not violate the law*). With this paraphrased translation together with the characters' facial expression, tone of voice, it is probably easier for Chinese viewers to appreciate the humorous meaning of the dialogue when watching this scene.

Source: Oh, you do, do you? Indeed, there isn't...

Translation: 哦? 你不会真的认为这样可爱吧? 揪头发确实没犯法。

Ò? nǐ bú huì zhēn de rèn wéi zhè yàng kě ài ba? jiū tóu fà què shí méi fàn fǎ.

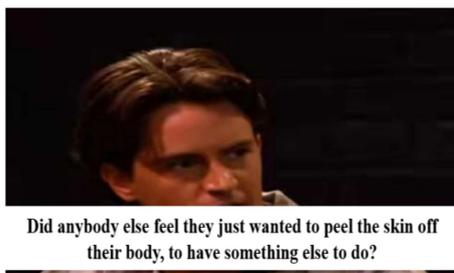


Figure 5. Example. 5 (Episode 6. 01:55).

In this scene, different discourses of the same auditory

modality form a joint relation and express a specific contextual meaning together. The context of this example is as follows: In the picture, the male character Joey's new musical show *Freud* was on stage, and his best friends were there to cheer him on. Before the performance, his friend Phoebe already implied that she thought the role of Freud was too difficult to play and Joey's performance was not good enough. The subtitle in the picture above is from another female friend, Monica who was commenting on Joey's performance. The implication of her comment *Did anybody else feel they just wanted to peel the skin off their body, to have something else to do?* was that if not for their friendship, she wouldn't have come to the show. What Monica said elicited a big laughter from viewers in the background, the two sounds of Monica's words and the background laughter, together with the sound of Joey's sadly organizing the props jointly express a strong ironical meaning of Monica's comment. In other words, different sources of voices, or different discourse of the same auditory modality, form a joint relation and jointly produce the ironical meaning. In translating Monica's words, the strategy of paraphrasing is used to give more explicit explanation of Monica's ironical intention and the subtitle is translated as 刚才有没有人宁愿剥皮拆骨, 或者做点儿任何别的什么事情, 也不想看这场演出? (*Would any rather have their skins peeled or do anything else but come to this terrible show?*). Such translation makes it clearer to Chinese viewers that Monica was commenting on a terrible performance. The translated subtitle that comes from paraphrasing, the sound of the props, and the laughter of viewers in the background combine in expressing the ironical meaning that Monica intended to convey.

Source: Did anybody else feel they just wanted to peel the skin off their body, to have something else to do?

Translation: 有没有人刚才宁愿剥皮拆骨, 或者做点儿任何别的什么事情, 也不想看这场演出?

yǒu méi yǒu rén gāng cái níng yuàn bāo pí chāi gǔ, huò zhě zuò diǎn ér rèn hé bié de shén me shì qing, yě bú xiǎng kàn zhè chǎng yǎn chū?

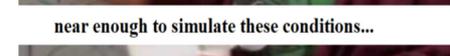
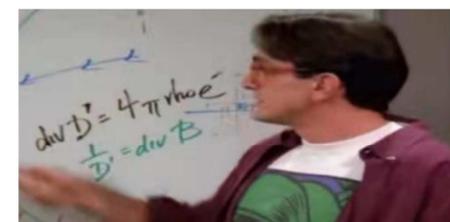
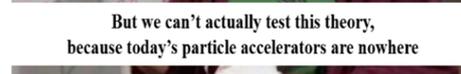
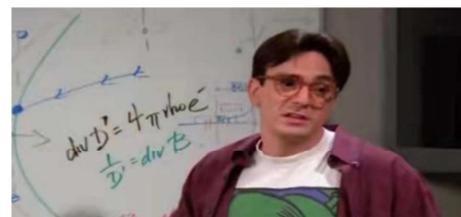


Figure 6. Example. 6 (Episode 10. 08:36).

In this example, discourses of visual and auditory modalities alternate to express the overall meaning. The context of this scene is as follows: It happened when the characters Phoebe and David met for the first time, and David was verbally explaining to Phoebe why it was impossible to test some theory in question (auditory discourse) and at the same time he was writing some complex physical equations on the whiteboard (visual modality). The visual modality and auditory modality in this case form an intersecting relation, and the meaning represented by either modality cannot fully express what the character David tried to convey, that is, it was still impossible to practice the theory in question. Taking this cross-modal relationship into consideration, this paper adopts the strategy of amplification in Chinese translation to include all the information that is crossmodally expressed in visual and auditory discourses. The subtitle *but you can't actually test this theory* in the first picture was translated as *但我们不可能实际去验证该理论* since the word *actually* was stressed by David in the auditory modality, an amplification of the phrase *实际去* (actually) transfers this paralinguistic stress into Chinese translation, keeping cross-modal meaning. The word *conditions* is translated as *方程式中所示的能量条件*, which refers to the energy of certain state shown in the equations in visual modality. Translation in this example integrates the information conveyed in both auditory modality and visual modality and the cross-modal information is verbally shown in translation, which allows Chinese viewers to accurately understand the meaning of the character.

Source: But you can't actually test this theory because today's particle accelerators are nowhere near powerful enough to stimulate these conditions.

Translation: 但我们不可能实际去验证该理论, 因为目前的量子加速器还不远无法达到创造方程式中所示的能量条件。

dàn wǒ men bú kě néng shí jì qù yàn zhèng gāi lǐ lùn, yīn wéi mù qián de liàng zǐ jiā sù qì hái yuǎn wú fǎ dá dào chuàng zào fāng chéng shì zhōng suǒ shì de néng liàng tiáo jiàn.

What is discussed above is the translation strategies used for multimodal discourses of non-reinforcing relation. In a multimodal discourse of non-reinforcing relation, discourses of two modalities complement each other and express an overall meaning together, and the translation strategy of amplification and paraphrasing makes the overall meaning displayed in two modalities stay complete as in Chinese translation in the form of single verbal modality.

4.2. Translation Strategies Regarding Multimodal Discourse of Non-complementary Relation

4.2.1. Translation Strategies Regarding Multimodal Discourse of Overlapping Relation



Figure 7. Example. 7 (Episode 8. 12:55).

In this scene, there is a redundant relation between two modalities. The scene is set after the death of the character Ross's grandmother. The character Rose was packing up his grandmother's belongings with his mother. Presented in the visual image is Ross finding a box containing a pink piece of paper that were beautiful sugar packets from his grandmother's collection. Hearing Ross's sigh, his mom asked what happened, but he only replied with a slight frown and a smile, "Yeah, just...just Nana stuff." In this example, what overlaps with what Ross said (auditory modality) is his expression and tone of voice, not the presentation of the pink piece of paper itself (part of visual modality). The presentation of the sugar packet does not provide additional meaning. Regarding the pink piece of paper, what viewers hear and see is the same referent, namely the sugar packet. Considering this redundant relation between visual and auditory modality, the strategy of direct translation is adopted in the Chinese translation and the subtitle is directly translated as *没事, 只是外婆的东西*.

Source: Yeah, just...just Nana stuff.

Translation: 没事, 只是外婆的东西。

méi shì, zhǐ shì wài pó de dōng xī.



Figure 8. Example. 8 (Episode 22. 17:38).

Sometimes, different discourses belonging to the same modality form an exclusionary relation against each other, which results in structural and semantic incompleteness. In *Example 8*, the character Phoebe, Chandler's secretary, found out that her colleagues did not like Chandler, and she told him truthfully that his employees would not like to be true friends with him as he was the boss, but Chandler disagreed and started defending himself. In auditory modality of this scene was a conversation between the two characters that has an obvious overlapping and interruption. According to the visual context, the full form of Chandler's words "I just want" would probably be "I just want to make friends with them (employees)", and the full form of Phoebe's words "but you can't" is "but you just can't make friends with them". Obviously, their conversation as auditory modal appears to be mutually exclusionary, forming an exclusionary relation making each other fail to complete their respective words. In Chinese translation, we adopt the strategy of amplification and translate both sentences into completed words to give Chinese viewers clearer information of the two subtitles that are incomplete in structure.

Sources: I just want...but you can't...

Translation: 我只是想和我的员工成为朋友; 但你作为老板是不可能和员工成为朋友的啊。

wǒ zhǐ shì xiǎng hé wǒ de yuán gōng chéng wéi péng yǒu; dàn nǐ zuò wéi lǎo bǎn shì bú kě néng hé yuán gōng chéng wéi péng yǒu de ā.



Figure 9. Example. 9 (Episode 10. 03:42; 04:16).

In this scene, there is a counteracting relation between two parts of auditory discourse. In detail, while Phoebe was singing a song with lyrics like *I made a man with eyes of coal and sometimes when it's freezing* etc., which is one part of auditory discourse in this case, two male customers were talking with each other off stage, one of whom said *I feel a little sneezy*, which is a secondary part of auditory discourse. Phoebe's singing was interrupted by two men's talk as is proven that later she became quite annoyed. According to Zhang [2], in this case two parts of auditory discourse form a counteracting relation in which Phoebe's singing was relatively louder than the two customers' talk. In our Chinese translation, the strategy of linear translation is used, which means to translate counteracting and overlapping film subtitles in lines, first louder and foregrounded sentences and then lower and backgrounded sentences and put translations of overlapping film subtitles on individual images. To be specific, the two lines are translated as follows:

Sources: I made a man with eyes of coal.

Translation: 我用煤当眼球堆了一个雪人。

wǒ yòng méi dāng yǎn qiú duī le yī gè xuě rén.

Sources: I feel a little sneezy.

Translation: 我有点想打喷嚏。

wǒ yǒu diǎn xiǎng dǎ pēn tì.

4.2.2. Translation Strategies Regarding Multimodal Discourse of Inclusive Relation



Figure 10. Example. 10 (Episode 5. 07:57).

In some cases, auditory modality and visual modality are in an abstract-concrete relation. In **Example 10**, characters Ross and Rachel planned to go to the laundromat to do their laundry together, but Ross's friend Chandler pointed out that this was undoubtedly a date for the two of them, and expressed his questioning on Ross bringing untidy underwear and a bottle of softener of brand *Snuggle*, to the laundromat for the date, because bringing the former would look rude and bringing the latter would make Ross immature as softeners of brand *Snuggle* is especially designed for children. Chandler's question to Ross, "*Oh, and uh, the fabric softener?*" is a grammatically incomplete sentence. Furthermore, although in their talking Ross held a bottle of softener and replied "*Ok, ok, now what is wrong with my Snuggles?*" in which the brand *Snuggles* was mentioned, it may still make Chinese viewers puzzled if Chandler's words are literally translated as 哦，那柔顺剂呢？(*Oh, how about the softener?*), which didn't explain why Chandler brought up softeners. The reference to softeners in auditory modality, that is, Chandler's words, is relatively abstract but the presentation of a bottle of softener of the specific brand of *Snuggles* in visual modality, although not producing new meaning, vividly display the reference. In other words, discourses of visual and auditory modality in this example are in an abstract-concrete relation. Considering this cross-modal feature and the need to explain to Chinese viewers why Chandler mentioned softener, the strategy of annotation and amplification is used. On one hand, an introductory line about *Snuggles*, 熊宝宝柔顺剂是美国家喻户晓的一款儿童衣物柔顺剂产品 is added as an annotation. On the other hand, Chandler's words "*Oh, and uh, the fabric softener?*" is translated into a grammatically complete sentence, 你不会还要带这个牌子的衣物柔顺剂吧？(*You aren't bringing this brand of fabric softener, are you?*). Ross's words "*Ok, ok, now what is wrong with my Snuggles?*" is translated as 行行行，但是我带熊宝宝牌柔顺剂有什么问题吗？(*What's wrong with me bringing Snuggles softener?*). with the help of annotation and amplification, information about the softener presented in visual modality is linguistically represented in Chinese translation, which better conveys Chandler's questioning that Ross was bringing a bottle of softener designed for children to a date.

Sources: *Snuggles*.

Translation: 熊宝宝柔顺剂是美国家喻户晓的一款儿童衣物柔顺剂产品。

xióng bǎo bǎo róu shùn jì shì měi guó jiā yù hù xiǎo de yī kuǎn ér tóng yī wù róu shùn jì chǎn pǐn.

Sources: *Oh, and uh, the fabric softener?*

Translation: 你不会还要带这个牌子的衣物柔顺剂吧?
 nǐ bú huì hái yào dài zhè gè pái zǐ de yī wù róu shùn jì ba?
Sources: *Ok, ok, now what is wrong with my Snuggles?*
Translation: 行行行, 但是我带熊宝宝牌柔顺剂有什么问题吗?
 xíng xíng xíng, dàn shì wǒ dài xióng bǎo bǎo pái róu shùn jì yǒu shén me wèn tí ma?

4.2.3. Translation Strategies Regarding Multimodal Discourse of Contextual-Interactional Relation

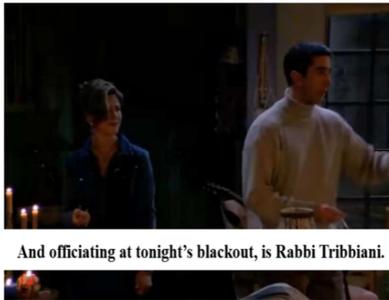


Figure 11. Example. 11 (Episode 7. 03:38).

In some cases, to understand what a character says needs to find correlated information in context shown in other modalities, under circumstance of which a dependent relation forms between discourse of different modalities involved. The word *blackout* in the character Ross's words "And officiating at tonight's blackout, is Rabbi Tribbiani." is associated with the dark atmosphere shown in the visual scene, but what the blackout refers to literally is not equivalent to what the visual scene shows, a party in a darkened room. In other words, auditory modality correlates with visual modality, but the latter provides additional information, and two modalities participate together in expressing the overall meaning in a dependent relation. Considering this cross-modal relation, when translating the subtitle, the strategy of free translation is used. Instead of directly translating *blackout* as 停电(power outage), it is translated as 不插电派对 (an unplugged party), which verbally represents all the cross-modal information in both visual and auditory modalities and facilitates Chinese viewers' understanding of the conversation.

Sources: And officiating at tonight's blackout, is Rabbi Tribbiani

Translation: 主持今晚不插电派对的是罗比·崔比雅尼。
 zhǔ chí jīn wǎn bú chā diàn pài duì de shì luó bǐ·cuī bǐ yǎ ní.

As is discussed above, when translating English subtitles into Chinese, the strategy of direct translation is used if the subtitle is in a redundant relation with discourse of other modality in the same scene; in an exclusionary cross-modal relation, the strategy of amplification can be adopted to give Chinese viewers clearer information of subtitles that are syntactically incomplete; in a counteracting relation, the strategy of linear translation is used to keep counteracting information verbally; for subtitles that are in inclusive relation, the strategies of annotation and amplification help to linguistically represent information and meaning that are cross-modally shown. As for subtitles that are in contextual-interactional relation, the

strategy of free translation is used.

5. Conclusion

This paper explores the translation strategies of English subtitles into Chinese from the perspective of multimodal interaction with the classical American TV show *Friend* as its corpus. Through the analysis of how discourses of cross-modal relations affect the translation practice, this paper finds that (1) For subtitles that belong to multimodal discourse of reinforcing complementary relation, it is advisable to make use of translational strategies of amplification, paraphrasing and annotation to highlight the foregrounded meaning expressed by the dominant modality on the one hand, and on the other hand to explicitly and verbally express the background information provided by the backgrounded modality if discourses of two modalities in the same scene form a relation of highlighting. However, it is more appropriate to make use of the strategy of free translation and paraphrasing to present the meaning expressed in the discourse of secondary modality when discourse of two modalities in the same scene are in the relation of primary-secondary relation. As for cases when discourse of two modalities belong to relation of expansion, the strategy of annotation can be adopted in translation. (2) For subtitles that belong to multimodal discourse of non-reinforcing complementary relation, discourses of two modalities complement each other and express an overall meaning together, and the translation strategy of amplification and paraphrasing makes the overall meaning displayed in two modalities stay complete as in Chinese translation in the form of single verbal modality. (3) For subtitles that is in multimodal discourse of non-complementary relation, the strategy of direct translation is used if the subtitle is in a redundant relation with discourse of other modality in the same scene; in an exclusionary cross-modal relation, the strategy of amplification can be adopted to give Chinese viewers clearer information of subtitles that are syntactically incomplete; in a counteracting relation, the strategy of linear translation is used to keep counteracting information verbally; for subtitles that are in inclusive relation, the strategies of annotation and amplification help to linguistically represent information and meaning that are cross-modally shown. As for subtitles that are in contextual-interactional relation, the strategy of free translation is used. In a word, it is both necessary and helpful to take cross-modal relations into consideration in English-to-Chinese subtitle translation in order to give Chinese viewers a better understanding of subtitles in context.

Despite the enlightening findings above, this paper specifically focuses on a qualitative analysis without quantitative consideration. Through collecting more data and building a larger corpus, a qualitative-quantitative study can be carried out in further research to testify the appropriateness and applicability of strategies of cross-modal translation that are drawn from case studies.

Acknowledgements

This research is a component of the 2018 research project

“Video-corpus-based Study of Chinese-speaking Children’s Multimodal Negating Act and Its Development” (18CYY022) sponsored by the National Social Science Foundation of China, the 2019 Shanxi Provincial Project of Philosophy and Social Science for Universities “A Video-corpus-driven Research on the Acquisition of Early-age Mandarin-speaking Children’s Multimodal Negation” (2019W001) sponsored by the Shanxi Provincial Education Department, the 2020 Shanxi Provincial Project of Philosophy and Social Science for Universities “A developmental study of multimodal negating of Chinese autistic children” (2020W021) sponsored by the Shanxi Provincial Education Department, 2021 Research Project “A developmental study of Chinese-speaking children’s multimodal referring” (2021-023) supported by Shanxi Scholarship Council of China, 2021 “A curriculum reform study of the core MTI selective courses under the guidance of Production-oriented Approach” (SXSkl2021SX0096) supported by Shanxi Federation of Social Science and Shanghai Foreign Language Education Press.

References

- [1] Christine, G. (2012). Reading the visual: Using Kress and van Leeuwen for a framework. *Metaphor*, 28-34.
- [2] Chen, HY. (2011). Multi-modal Subtitle Translation: A Case Study of *Gone With the Wind*. *Film Review*, 19, 71-73.
- [3] Guo, JZ. (2009). Foreignization and Domestication: Moral Attitude and Discourse Strategy-A Review of the Second Edition of Venuti’s *Translator’s Invisibility*. *Chinese Translators Journal*, 194 (2): 34-38.
- [4] Hu, MX, & Li, ZZ. (2018). Multimodal Speech: Conceptual Analysis and Research Prospects. *Journal of PLA Foreign Languages University*, 41 (03): 10-18.
- [5] Kan, F. (2014). *Film Subtitle Translation from the Perspective of Multi-modal Discourse Analysis*. Qiqihar University.
- [6] Liu, CL. (2016). An Interpretation of Domestication and Foreignization Strategies in Film and Television Subtitle Translation-A Case Study of *Confessions of a Shopaholic and Friends*. *Journal of Hunan First Normal University*, 16 (6): 91-95.
- [7] Li, XT. (2019). Research on Multimodal Interaction and Chinese Multimodal Interaction. *Language Teaching and Research*, 4, 45-49.
- [8] Li, YX. (2001). The Strategies of Subtitle Translation. *Chinese Translators Journal*, 4, 38-40.
- [9] Lv, JY. (2008). Strategies of Film Subtitle Translation-Taking English Film Translation as an example. *Film Literature*, 5, 109-110.
- [10] Wang, RB & Zhang, JY. (2020). An Empirical Study of the Development of 4-8-year-old Children’s Verbal-gestural Acts. *Foreign Language Education*, 41 (4): 69-74.
- [11] Wei, QH. (2008). On the Overall Meaning Construction of Multimodal Discourse-Discourse Analysis Based on Multimodal Media Discourse. *Journal of Tianjin Foreign Languages University*, 6, 16-21.
- [12] Xu, MJ. (2017). A Review of Multi-modal Translation Studies in China. *Journal of Guangdong University of Foreign Studies*, 28 (2): 40-46.
- [13] Yuan, FL. (2010). Strategies for Translating Subtitles into Chinese-A Case Study of the American Sitcom *Friends*. *Journal of Tianshui Normal University*, 30 (1): 126-128.
- [14] Zhang, DL. (2009). Research on Comprehensive Theoretical Framework of Multimodal Discourse Analysis. *Chinese Foreign Languages*, 6 (1): 24-30.
- [15] Zhang, DL. (2018). A Comprehensive Framework for Multimodal Discourse Analysis from the Perspective of Systemic Functional Theory. *Modern Foreign Languages*, 41 (6): 731-743.