On the Multimodal Analysis Model of Subtitle Translation: A Case Study of *Wolf Totem*

Yan Qian[1], Tian Chuanmao[2]

[1]School of Foreign Studies, Yangtze University, Hubei, 434023 PRC China
Email: 1449113309@qq.com

[2]School of Foreign Studies, Yangtze University, Hubei, 434023 PRC China
Email: tcm_316@163.com (Corresponding Author)

Abstract— *Film is considered as a visualized and vivid multimodal communication mode, the cross-cultural dissemination of which cannot be separated from its subtitle translation. Based on the Synthetic Theoretical Framework of Multimodal Discourse Analysis (STF of MDA) put forward by Professor Zhang Delu, the paper analyzed the characteristics and effects of subtitle translation in Wolf Totem from four levels, namely culture, context, content and expression, exploring how different modalities coordinated to construct the whole meaning of the film discourse. After detailed analysis on typical examples, it raised a multimodal analysis model of subtitle translation on the basis of Zhang's theory, so as to provide a new analytical approach for research on subtitle translation.*

Keywords— *Wolf Totem; subtitle translation; multimodal discourse analysis; multimodal analysis model.*

I. INTRODUCTION

In the recent dozen years, multimodal analysis of subtitle translation has received much attention from domestic and foreign scholars, some of whom explored the influence of non-verbal symbols on subtitle translation. However, most of the studies focused on the effects of multimodal elements in static frames but ignored the dynamic nature of film discourse (Chuang, 2006; Pérez-González, 2014; Silvester, 2018). Moreover, some scholars made multimodal analysis of subtitle translation form the perspective of varied theories. For example, Taylor (2003) applied the Multimodal Transcription Theory to investigate how utterance interacts with other meaning-construction semiotic resources by means of subtitles. Based on STF of MDA, Lv & Wu analyzed subtitle translation of *The Flowers of War* from four levels of culture, context, content and expression. Chen & Wang (2016) studied the key roles played by dynamic frames in subtitle translation from three

aspects of Visual Grammar: representational meaning, interactive meaning and compositional meaning. These researches made on the basis of different theories certainly enriched the research content of subtitle translation. What is particularly worth mentioning is that there are foreign scholars who studied how to translate auditive information into verbal one in target language under the influence of frames. This kind of research could help to solve the problem that hearing-impaired people of target language are unable to receive auditive information effectively (Hurtado & Martínez, 2018). Generally speaking, most previous studies at home stay at the level of case analysis, and there are merely two related theoretical researches. Chen & Zhang (2017) built up a theoretical analysis model for subtitle translation according to the multimodal analysis method under the framework of Systemic Functional Linguistics. Lv (2016) integrated visual modality and auditive modality into the meaning-generating mechanism of subtitle translation and then

constructed multimodal contextual parameters. We hold that multimodal translation study of film discourse is in urgent need of the guidance of systemic theoretical model and analytical framework, and thus future related studies had better be devoted to theoretical integration and innovation.

The film *Wolf Totem* is in the shortlist of Tiantan Awards of the 5th Beijing International Film Festival, having won Best Director Award and Best Visual Effect Award. In the film, subtitles coordinated with frames, sound and other symbols closely to construct the whole meaning of the film discourse, which created and presented favorable visual and appreciating effects. The paper aims

to study translation features and effects of the film *Wolf Totem* from culture, context, content and expression based on STF of MDA, and attempts to integrate related theories to construct a multimodal analysis model for subtitle translation.

II. SYNTHETIC THEORETICAL FRAMEWORK of MULTIMODAL DISCOURSE ANALYSIS

Drawn from Systemic Functional Linguistics of Halliday, Zhang Delu (2009) put forward Synthetic Theoretical Framework of Multimodal Discourse Analysis (STF of MDA) mainly consisting of four levels: culture, context, content and expression.
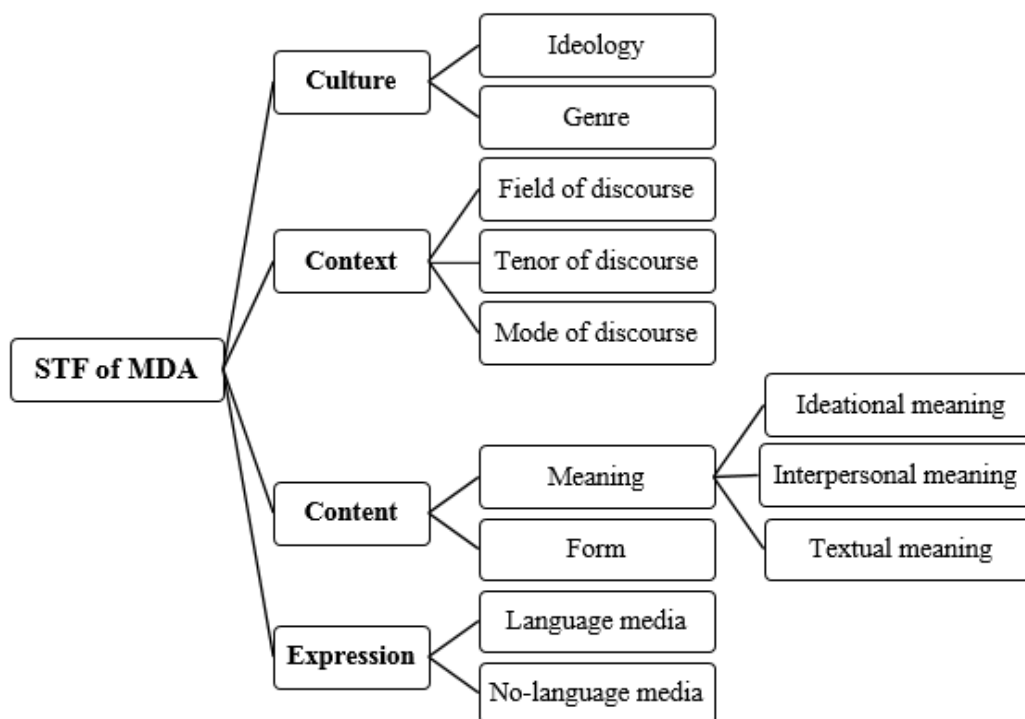

Fig.1: Synthetic Theoretical Framework of Multimodal Discourse Analysis

(STF of MDA) (Zhang, 2009)

The culture level is the key to make multimodal communication possible, which determines the pattern, form and skills of communication. Without the culture level as foundation, context of situation will lose its explanatory power. The culture level includes ideology made up of human's thinking pattern, philosophy of life, living habits

and all hidden social rules, and genre which is a communicative procedure to realize the ideology. In specific contexts, communication would be restricted by field of discourse, tenor of discourse and mode of discourse. The content level could be categorized into semantic level that is composed of ideational meaning, interpersonal meaning and

textual meaning and form level. In terms of form level, every modality owns its form system, like visual grammar, auditive grammar and tactual grammar, and varied modalities coordinate with each other to achieve the best communication effect. The expression level is divided into language media and non-language media. From the perspective of traditional linguistics research, language media refers to sound symbol conducted by acoustic wave and written symbol generated by writing. While non-language media conveying meanings contain the postures of communicator and non-verbal means used by communicator like tools and environment (Zhang Delu, 2009).

In addition to STF of MDA, Visual Grammar (Kress & Van Leeuwen, 1996) and Framework of Intersemiotic Complementarity (Royce et al., 2007) are often adopted to study film subtitles translation. Nevertheless, STF of MDA is believed to have stronger explanatory power and applicability for subtitle translation. Visual Grammar attaches undue importance to meanings conveyed by frames, but ignores the coordination between frame and words. Although Framework of Intersemiotic Complementarity makes up for its deficiency and emphasizes the complementary relationship between frame and words of visual modality, film discourse is also involved in auditive modality. Generally, Zhang Delu's STF of MDA fully considers symbols of visual modality like frame, words, movement and facial expression, and symbols of auditive modality like sound and tone, all of which are responsible for constructing and conveying meanings, so it's more applicable to multimodal analysis of subtitle translation.

III. ANALYSIS ON SUBTITLE TRANSLATION OF WOLF TOTEM FROM THE MULTIMODAL PERSPECTIVE

Wolf Totem is an adventure film made by China and France, adapted from Jiang Rong's novel with the same name, which is one of the classic works of the notable French director Jean-Jacques Annaud. The story happening in 1967, two school graduates of Beijing Chen Zhen and his friend Yang Ke went to the grassland of Inner Mongolia living and working in a production team. On the broad

grassland, they met the Mongoloid herdsman Bilig and his family members, and they really knew the most awful animal of the grassland—wolf. During his contact with wolf pack, Chen Zhen was deeply fascinated by wolves and showed keen interest in the species. He even captured a wolf cub and fed it at home, so as to explore the habits and characteristics of wolf. At that moment, a group of farming people from the eastern part of Inner Mongolia greedily plundered the gazelles stored by wolves for winter. Moreover, the head of the production team Bao Shungui organized an activity of ridding the wolves' lairs of their cubs. The wolf pack was entirely infuriated by the two merciless behaviors of human beings. In order to take vengeance on greedy ruthless humans, the wolf pack made use of blizzard of winter and mosquito disaster of summer as their shields, and finally succeeded in raiding and killing horse pack and gazelle pack. Farming people were absolutely furious at wolves, and thus they took extreme measures like burying explosives and setting on fire. Consequently, wolves on the grassland were all annihilated. The last untapped grassland and the last clean lake of Inner Mongolia were destroyed. At the end of the film, the chief actor Chen Zhen was leaving the grassland, the released wolfling loomed and then disappeared on the withered and yellow grassland.

The authors are going to discuss four research questions from four levels of STF of MDA, culture, context, content and expression, combining typical examples of *Wolf Totem*: (1) the reinforcement relationship between frame and subtitle & analysis of culture-loaded words; (2) the complementary relationship among sound, frame and subtitle & analysis of situational meaning; (3) ternary interplay among sound, frame and subtitle & analysis of discourse meaning; (4) typesetting of subtitle of language media & analysis of visual grammar and auditive grammar of non-language media.

3.1 The Culture Level

As an intuitionistic way of cultural dissemination, film bears abundant cultural connotations. Subtitle translation is not only transformation at the language level, but also delivers cultural information, because language carries cultural genes (Tian & Wang, 2017). Therefore, translators

are supposed to have cultural consciousness and realize that subtitle translation is a communication process cross language and culture, making efforts to overcome barriers caused by cultural differences and ensure smooth realization of emotional information exchange. (SL—source language; TL—target language).

Example 1:	
SL	那是 1967 年，文化革命的第二年，城里的学生被下放到农村去插队。
TL	That was in 1967, the second year of the Cultural Revolution. City students were sent to the countryside.

Example 1 is extracted from the beginning of the film. The monologue of chief actor Chen Zhen introduced historical background of the story, establishing tenor of discourse for the development of plots. The word "插队" refers to the pattern that educated urban youth went to work in the countryside and mountainous areas before 1980. In another word, the youth in urban areas were sent to rural production team to engage in agriculture. That is completely different from the meaning in our daily life where "插队" is used to indicate "jumping a queue". The subtitle translator didn't translate the culture-loaded word "插队" and translated "下放" into "sent", the cultural coloring of which is blurry. However, the cultural context implied by the sensitive culture-loaded word "文化革命" (Cultural Revolution) and the film frame that many educated youth went to Inner Mongolia grassland by bus to receive reeducation from the poor and lower-middle peasants enable foreign audiences to understand original meanings and background of the story accurately.

Example 2:	
SL	伟大的腾格里，是他唯一的永恒的主宰，他只听从腾格里的号令。
TL	He would only answer to Great Tengger，his one and only Master.

In the folk religion of Mongoloid, "Tengri" is supreme god regarded as the dominator of world and human beings, the Mongol of which is "Mongke Tengri" (referring to the entry "腾格里" in Baidu baike). In example 2, the translator adopted the strategy of foreignization to transliterate the word with Chinese cultural characteristics into "Tengger" from the perspective of Mongol, which retains the unique religious culture of Mongoloid to the greatest extent. For those who are not familiar with Chinese culture, it's difficult for them to understand the actual meaning of "Tengger", but the word "Master" after it could make target audiences enlightened who will know Tengger is the dominator of everything. In addition, the film frame that farming people throw wolflings to the sky hoping they go to Tengger to get immortality appears several times. Under the help of those frames, target audiences are able to understand the translated subtitles. The example used the method of reordering to cater the expression habits of English, which could be considered as domestication at the language level. The translator considered the influence of cultural factors on subtitle translation and adopted translation strategies of foreignization and domestication appropriately. As a result, culture-loaded words were translated in a concise and clear way, having ensured smooth understanding of target audiences without losing cultural characteristics of original words.

3.2 The Context Level

The relationship between modality and context could be regarded as that between positive modality and negative modality (Zhang, 2009). Context falls into textual context, situation context and cultural context (Hu, 2007), and situation context is a significant factor affecting subtitle translation. Translators could take full advantage of contextual information and add or delete verbal messages appropriately in order to enable frames to be presented naturally and smoothly.

Example 3:	
SL	再多攒点钱就买得起了
TL	Bring us more gazelles then

The herdsman Shartseren sold the gazelles he had picked up in the snow lake and wanted to buy a radio from the buffet. The price of the radio was 120 RMB, but there

was only 80 RMB in his pocket. The word "钱" was translated into "gazelles" instead of "money", because the translator had fully considered the previous plots that Shartseren exchanged gazelles for money and that farming people of the eastern part endeavored to inquire about the location of the snow lake. The translated subtitle indicates that selling gazelles was an important way of making money for those farming people and is also in accordance with the following plot that eastern farming people went to the snow lake and plundered all the gazelles stored by wolves for winter.

Example 4:	
SL	马群里有狼
TL	Wolves

On a night of snowstorm, the wolf pack took vengeance on humans and attacked the horse pack on a large scale. Batu and Shartseren tried to control the horse pack, but they desperately found that wolves had sneaked into the horse pack. In example 4, the translator used the skill of omission and just translated the word "狼" with the adverbial "马群里" deleted, which accords with the story context. From the frames, target audiences could clearly obtain the information that there were wolves in the horse pack. Moreover, the background music was nervous and hurried and target audiences paid much more attention to screens, so it's inappropriate for the translated subtitle to be too long. We can say that the translator got the utmost out of contextual information to conduct suitable deletion.

Example 5:	
SL	外边草原上的野狼才是你们的真正的敌人呢。我这次来，就是帮着你们消灭敌人了。
TL	It is the wild wolves that are your problems. And I am here. With the solution.

The wolf pack killed all the sheep fed by local farming people and there was no one left. The farming people was filled with anger and owed it to the wolfling fed by Chen Zhen. From their point of view, it was the wolfling that had incurred the visit of wolves to their populated areas. Out of

this reason, they planned to kill the wolfling to vent their resentment, but Chen Zhen was fighting tooth and nail. At the moment, the leader of the production team Bao Shungui showed up and prevented the fight and quarrel. In example 5, the sentence "就是帮着你们消灭敌人了" was translated into "With the solution" briefly, the translator having considered the restrictions on time and space of subtitles. And the translated subtitle worked in concert with the following plot that Bao Shungui guided his production team to catch and kill wolves on the grassland by pouring gasoline and gun killing.

3.3 The Content Level

In the respect of semantic meaning, translators usually make use of the flop-out method where interpersonal meaning and discourse meaning are cut down as a general rule because of the restrictions on time and space of subtitles (Yang, 2012). When it comes to subtitle translation of *Wolf Totem*, we found that modal particles and address forms were usually omitted because these words would not affect the delivery of whole meanings of film discourse and target audiences could understand original meanings under the help of frames and background sound effects.

Example 6:	
SL	陈阵，你啥时候也和蒙古族的女人，生个娃，那才能称得上真正的蒙古人呢。
TL	When do you have a Mongolian woman and a baby，that can be called the real Mongols.

Example 6 is extracted from the frame that a group of Mongolian herdsmen were sitting on the grassland and someone was playing tricks on Chen Zhen and Gasma. In the translated subtitle, the address form "陈阵" and the modal particle "呢" were cut out, but it would not influence target audiences' comprehension of original meanings at all. From the frame, they could know the tricked object was Chen Zhen and that the joking tone of the speaker was evident due to relaxing background music and merry facial expressions.

During the process of subtitle translation, visual modality and auditive modality will get involved pure

language symbol, and the three modalities interplay to construct whole meanings (Lv, 2016). In film discourse, the form level mainly shows as utterance, background music and other sound effects of auditive modality, and frame and subtitle of visual modality. The complementation and reinforcement of varied symbols endow a film with more attractive power. For instance, the melodious cantus of Morin khuur that is an instrument with Mongolian folk characteristics rose at the moment of 2 minutes and 38 seconds, the background sound effect made audiences relaxing. And the cool horse riding performance of Mongolian herdsmen present an atmosphere of jubilation and jollification. At 42 minutes and 15 seconds, when wolves attacked the horse pack at a snowstorm night, the frames that the panic horse pack were out of control and wolves were chasing after horses impetuously interlaced with nervous music, clip-clop, howl, shouting of humans and corresponding subtitles, which created favorable visual and auditive effects and made it much easier for target audiences to understand film plots.

3.4 The Expression Level

In terms of language media, film conveys meanings through subtitles which play a significant role in the whole expression of meanings. A translated subtitle should finish information delivery in roughly the same time as that of source language and each subtitle should stay on the screen with enough time for audiences to browse, generally 2 to 3 seconds (Li, 2001). In view of the limitation of screen space, an English subtitle of film are supposed not to exceed two lines and a Chinese subtitle should be controlled within one line (Asher & Simpson, 1994). There are bilingual subtitles at the bottom of screen in *Wolf Totem* where the top line is Chinese subtitle and the bottom line is English subtitle. Bilingual subtitles of both Chinese and English could help audiences of different types to obtain information they need respectively so as to understand the whole film discourse. Generally speaking, film subtitle is monologue or dialogue of characters, but we hold that it's necessary for translators to add subtitles for critical non-verbal messages.

Example 7:	
SL	热烈欢迎知识青年上山下乡来内蒙古安家落户
TL	Welcome educated youth to Inner Mongolia and get settled.

At the moment of 1 minute and 50 seconds, as soon as Chen Zhen, Yang Ke and other urban educated youth came to Inner Mongolia, there was a red streamer hanging on the wall, which read "热烈欢迎知识青年上山下乡来内蒙古安家落户". There was a close-up shot for the slogan that told audiences the background and location of the story, but the non-verbal information wasn't translated. As a matter of fact, audiences of target language is not familiar with Chinese historical background during that time. Therefore, an English subtitle for the slogan on the screen will be beneficial to foreign audiences' comprehension of film plot and related cultural backgrounds.

In addition to language media, translators also need to pay attention to non-language media generating meanings. Non-language media mainly includes bodily media like movement of communicator, and non-bodily media like tools used by communicator to convey meanings or external environment (Zhang, 2009). At the moment of 32 minutes and 58 seconds, the wolflings gathered by Chen Zhen and Yang Ke with much efforts were thrown to the sky by a merciless herdsman. Chen Zhen was astonished and desperate, he sat on the ground weak and limp, hugging the gunny bag used to contain wolflings and fondling it gently, the frowned brows revealing his sadness and anger. Yang Ke stretched out his hands slowly to touch the dead wolflings, but he was shoved down the ground by Chen Zhen because he didn't prevent the herdsman from killing their wolflings. All of these physical movement are bodily media. While non-bodily media include the blood-stained gunny bag, overcast sky and withered and yellow grassland. The frame took full advantage of non-language media to heighten a dismal atmosphere.

IV. MULTIMODAL ANALYSIS MODEL OF SUBTITLE TRANSLATION

Previous case analysis proves that Zhang's STF of MDA is feasible for multimodal analysis of subtitle translation. Film subtitle translation is not simply the transformation of verbal symbols, and varied non-verbal symbols are involved in the translation process, the influence of which cannot be ignored. In order to highlight the joint efforts of verbal and non-verbal symbols in visual modality and auditive modality, the authors integrated Zhang's framework, Visual Grammar and Systemic Functional Grammar to construct a multimodal analysis model of subtitle translation (Figure 2).
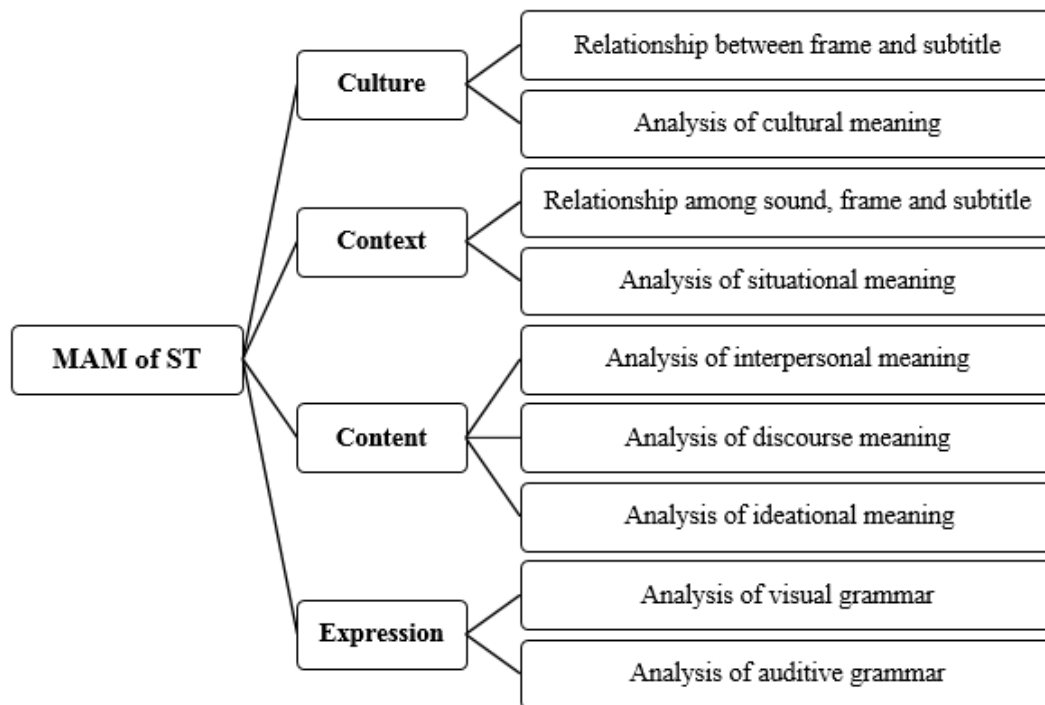

Fig.2：*Multimodal Analysis Model of Subtitle Translation (MAM of ST)*

Comprehensive evaluation of the coordinating effects of varied modalities could be made from four aspects: culture, context, content and expression. At the culture level, we could attach importance to the reinforcement relationship between frame and subtitle, and the connation of culture-loaded words, investigating the complementary role played by the frame symbol to supplement background information. Appropriate and effective means of disseminating culture like omission could help audiences of target culture appreciate and comprehend original plots in uninterrupted context of frames. At the context level, the complementary relationship among sound, frame and subtitle and situational meaning could be analyzed. Taking textual context and situational context created by current frame and sound into full consideration, translators might adopt translation skills like substitution, amplification and omission to integrate translated subtitles with sound and frame so that the best visual effect could be achieved. At the content level, the interaction among sound, frame and subtitle and semantic meaning could be discussed, including interpersonal meaning, discourse meaning and ideational meaning. In consideration of the fact that mood and tone of dialogue are the important parts of emotional meanings, we are supposed to be prudent to omit modal particles when translating subtitles. At the expression level, subtitle typesetting of language media should be paid attention to, moreover, it's necessary and significant to explore how frames, visual grammar and additive grammar of non-

language media realize the generation of specific meanings. For instance, how body language of characters, props and external environment cooperate with background music to create a situation that cannot be achieved through a single modality.

V.   CONCLUSION

Under the guidance of Zhang's STF of MDA, the paper discussed the subtitle translation of *Wolf Totem* from culture, context, content and expression, on the basis of which a multimodal analysis model of subtitle translation was proposed. The research found that there are at least reinforcement relationship and complementary relationship between translated subtitles and non-verbal symbols. Words, frames, sound effect, background music and other semiotic resources coordinate with each other to construct the whole meaning of film discourse. In another word, subtitle translation is different form the transformation of single language modality. Only when translators take the particular roles played by non-verbal symbols into consideration, could translators produce subtitles of target language closest to plots. Based on multimodal analysis of subtitle translation, we believe that it's unnecessary and unwise to translate subtitles words by words, and that only core information needs to be translated to achieve reliability and validity of subtitles because audiences could obtain other subordinate information from the involvement of non-verbal modalities. Additionally, it's necessary for translators to add corresponding subtitle of target language for non-verbal information that might influence the understanding of target audiences.

REFERENCES

[1] Asher, R. E., & Simpson, J. M. Y. *The encyclopedia of language and linguistics* [M]. Oxford: Perganon Press, 1994.

[2] Chen, Y., & Wang, W. Relating visual images to subtitle translation in Finding Nemo: A multi-semiotic interplay [J]. *Translation & Interpreting*, 2016, 8(1): 69-85.

[3] Chen, Y. P., Zhang, C. H. On Chinese translation of English film subtitles—A multimodal analysis of the relationship between images and words [J]. *Chinese Translators Journal*, 2017(5):105-110.

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

[5] Hu, Z. L. Multimodality in socio-semiotic studies [J]. Language Teaching and Linguistic Studies, 2007(1):1-10.

[6] Hurtado, C. J., & Martínez, S. M. Concept selection and translation strategy: Subtitling for the deaf and hard-of-hearing [J]. *Linguistica Antverpiensia, New Series: Themes in Translation Studies*, 2018(14): 114–139.

[7] Kress, G. R., & Van Leeuwen, T. *Reading images: The grammar of visual design* [M]. London: Psychology Press, 1996.

[8] Li, Y. X. Translation strategies of subtitles [J]. *Chinese Translators Journal*, 2001(4):38-40.

[9] Lv, J., Wu, W. Z. On subtitle translation of *The Flowers of War* based on Multimodal Discourse Analysis [J]. *Shanghai Journal of Translators*, 2012(4):36-38.

[10] Lv, J. The meaning-generating mechanism of subtitle translation under MCPT: A case analysis of *Blood and Bone* [J]. *Foreign Language and Literature*, 2016, 32(6): 128-135.

[11] Pérez-González, L. *Audiovisual translation: Theories, methods and issues* [M]. London: Routledge, 2014.

[12] Royce, T., Royce, T. D., & Bowcher, W. L. Intersemiotic complementarity: A framework for multimodal [J]. *New Directions in the Analysis of Multimodal Discourse*, 2007: 63-109.

[13] Silvester, H. From paratext to polysemiotic network: A holistic approach to the study of subtitled films [J]. *Linguistica Antverpiensia, New Series: Themes in Translation Studies*, 2018(14): 71–83.

[14] Taylor, C. J. Multimodal transcription in the analysis, translation and subtitling of Italian films [J]. *The Translator*, 2003, 9(2): 191–205.

[15] Tian, C. M., Wang, F. Translation and Culture [M]. Beijing: China Social Sciences Press, 2017.

[16] Yang, W. W. On subtitle translation of Gone with the Wind from the perspective of multimodal discourse analysis [J]. Movie Literature, 2012(19):155-156.

[17] Zhang, D. L. Synthetic Theoretical Framework of Multimodal Discourse Analysis [J]. Foreign Languages in China, 2009(1): 24-30.