A Study of Subtitle Translation from the Perspective of Multimodal Discourse Analysis: A Case of *My People, My Country*

Minggui Duan

College of Foreign Languages, Shanghai Maritime University, Shanghai, China
richardduan@foxmail.com

Abstract: *In the context of globalization, audio-visual works produced by ever-evolving information technology provide a quick and convenient channel for bridging cultural divides and communicating with people from all over the world. People harness these productions not only to satisfy their own entertainment needs but also to gain information about other cultures. As manifold cultures and races become increasingly interconnected, the need for intercultural communication is particularly urgent, and with abundant European and American films and cultural productions taking over the world, the significance of subtitle translation cannot be understated. On the surface, subtitle translation transforms one language into another and puts it on screen at the same time. However, at a deeper level, subtitling involves a wide range of cultural and social contexts in both the source and target languages. Translators need to not only use their translation skills to select words and sentences, but also have a deep understanding of the social and cultural aspects of both languages. This is the only way to smoothly accomplish the language transfer and help the audience accurately make sense of the meaning of the words in the film. At the same time, China has entered a new era and is facing unprecedented changes in a century. Only by spreading its own culture, promoting the achievements of excellent Chinese civilization, and gaining widespread international recognition will it be able to stand firm and progress against the tide of globalization. However, even though Chinese film and television products are constantly being produced and innovated every year, there are certain shortcomings in the dissemination of film and television works and the study of subtitle translation. This thesis takes Professor Zhang Delu's theoretical framework of multimodal discourse analysis as its perspective and selects the subtitle corpus of the film My People, My Country, which was released around National Day in 2019 to celebrate the 70th anniversary of the founding of the People's Republic of China, as the object of study. Based on the integrated framework of multimodal discourse analysis, the author carefully selects typical examples from My People, My Country to analyze the effects of multimodal discourse analysis on subtitle translation at the cultural, contextual, content, and expression levels, and investigates how different modalities work together to vividly convey the meaning of the film's discourse. In conclusion, multimodal discourse analysis theory has injected new momentum into the study of subtitle translation applications, breaking away from the traditional monomodal and symbolic studies and contributing to the international dissemination of more domestic audio-visual cultural products.*

Keywords: Multimodal discourse analysis, Subtitle translation, *My People, My Country*.

1. Introduction

As a kind of intermediary, which transmits what the audio-visual product originally means to the target audience who would like to make sense of such meanings, subtitle translation calls for professional treatment to ensure that people from other cultural backgrounds can attain the maximum and most faithful information from audio-visual products like films and TV series. To some extent, whether the subtitle translation is treated appropriately or not, has a tremendous impact on the target audience's understanding of the product, especially in the activity of cross-cultural communication. On that score, an apt theoretical approach to the arrangement of subtitle translation is of great significance.

At present, some translation theories are not appropriate in research in audio-visual subtitle translation. In this case, a more desirable theory will greatly benefit researchers by conferring upon them an advanced strategy for audio-visual subtitle translation. At the same time, a more satisfying viewing experience will be attained by the target viewers. Breaking out of the conventional merely focus on texts, multimodal discourse is an integration of verbal and nonverbal communicative modes embodying language, images, sounds, and so on (Kress & Van Leeuwen, 1996). Multimodal discourse analysis derives from linguistics to delve into semiotic meanings, which is often exerted in films or TV series. It is conceivable that multimodal discourse analysis will function as a unique conception, invigorating academic research on subtitle translation.

Under the arrangement of multimodal discourse analysis on the subtitle translation of the film *My People, My Country*, it can be uncovered what the effects are and how they are presented in the film. As well, subtitle translation will be treated in a different way, which is favorable to research concerned in the future and can exploit a new spectrum in cross-cultural communication study.

Generally, under the guidance of the theoretical framework of multimodal discourse analysis proposed by functional linguist Zhang Delu, the paper presents a relatively systematical analysis of the subtitle translation of the film—*My People, My Country*, which was released for celebrating the 70th founding anniversary of People's Republic of China. Specifically, this paper is essaying to investigate whether the translation version of the film was rendered in consistence with the four respects of cultural level, contextual level, content level and expression level so that it can achieve spectacular effects by firstly ensuring that the audience can make a better comprehension of the connotations or intentions revealed by the film. Furthermore, this paper is also considering to discover different modalities demonstrated in this film with the following modalities to be explored—sound, images, videos and so on. Correspondingly, this paper will be conducted by delving into these following questions:

1) By applying the techniques of the theoretical framework of multimodal discourse analysis, how does the subtitle translation of *My People, My Country* propel a deeper understanding of foreign viewers?

2) What effects will this subtitle translation of *My People, My Country* bring to us from the perspective of multimodal discourse analysis?

3) How do different modalities in the film *My People, My Country* interact with each other for the sake of information transmission and cultural communication?

Through purposive and rational responses to these three questions in the process of the research, this paper will promisingly accomplish its goals, attaining practicable and scientific results. At the meantime, the thesis is also aimed at drawing a conclusion that the application of multimodal discourse analysis into the subtitle of *My People, My Country* is an advantageous tactic for information dissemination of film media. Finally, it is expected that by using effective translation methods, Chinese cultural works can be appreciated and embraced by more and more international viewers.

2. Literature Review

Based on functional linguistics theory and cognitive linguistics, multimodal discourse analysis has bestowed upon us a new sight into the theory on analytical study on manifold discourses produced during various activities existing in human society. In recent decades, many researchers have increasingly shifted their interests into the research on multimedia discourses like subtitle translation or images by applying the theory of multimodal discourse analysis. A detailed and informative summary is of great significance to us for it can not only present related achievements to us but also throw light on the previous and current explorations conducted by those who have delved into this specific domain.

2.1 Previous Studies on Multimodal Discourse Analysis

2.1.1 Studies Abroad

Since American linguist Harris (1952) proposed the theory of discourse analysis, researchers at home and abroad have been engaged with the construction of different discourse analyses both theoretically and methodologically. In the 1990s, the study of multimodal discourse analysis began to thrive in the international linguistic spectrum. O'Toole (1994) dissected the practical analysis of multimodal discourse by investigating the visual works of art from the grammatical angle. He elaborated on the concept of model meaning when people view a visual work of art. Linguists Kress & Van Leeuwen established the theoretical generalization of multimodal discourse analysis. They described the term multimodal discourse as "a phenomenon of communicating through manifold means such as languages, images, sounds, actions and symbolic systems by using senses like hearing, vision, touching, etc.." (Kress & Van Leeuwen,1996)

In the 20th century, the theory of multimodal discourse analysis is subsequently demonstrated in a case study of modern multicultural products by academics' deep exploring the existing theoretical basis and expanding it into the aspect of practical analyses. Kress & Van Leeuwen made an expansion of the meaning of multimodality, defining multimodality as the "use of several semiotic modes in the design of a semiotic product or event" (Kress & Van Leeuwen, 2001, p. 20-22). Multimodal analyst Norris (2004) furthered enriching the theoretical body of multimodal discourse analysis by putting forward an integrated methodological framework consisting of a combination of verbal and non-verbal when in communication. John, Bateman & Janina (2014) pointed out a pragmatic model of multimodal discourse resting on linguistics and provided a comprehensive methodology for studying multimodal artifacts.

2.1.2 Studies at Home

Li Zhanzi was the first to introduce this framework into the Chinese academic field of linguistics and acted as a spearhead in the exploration of a new branch of linguistics.

Afterwards, domestic experts have theoretically carried out a series of related researches, which has greatly enriched different notions of multimodal discourse analysis. Li (2003) has illustrated the methods of visual syntax and images analysis built up by Kress & Van Leeuwen. She investigated the frame of images analysis from the perspective of sociology, analyzing the meanings conveyed by the syntactic relations among people, places and events seen in images. After this, Hu (2007) pointed out that semiotics should focus on more extensive objects, namely, senses like seeing, listening, smelling, tasting or even abstract models, constructions and subjective consciousness should be put into the field of multimodal discourse research. Zhu (2007) made discussions about the theoretical foundation as well as the approaches to the research concerned. Zhang (2009) focused on the exploration of the theoretical framework of multimodal discourse analysis. In the meantime, he expounded on the mutual relationship of different multimodalities between each other. In addition, the establishment of the grammatical structure of nonverbal modes, the relationship between grammatical structure and textual structure, and how to realize the functional meaning of grammatical structure have also been delved into by him. Wei (2009) analyzed the discrimination between mode, medium and modality.

Then, in recent ten years, scholars have furthered their research on multimodal discourse analysis from more different angles. Hu (2014) explored the translatability of multimodal discourse from the perspective of textual typology. He held that by choosing different translation principles and strategies according to different types of texts, a successful expression of the true intention of the original text can thus be made. Yu (2018) criticized the hypothesis that translation texts are abstract objects constructed by linguistic symbols and advocated studying the visual text from the perspective of multimodal discourse analysis. Zhang & Zhao (2021) expounded on the models which constitute multimodal discourse and then offered professional devices to analyze multimodal discourse at the grammatical level.

2.2 Previous Studies on the Subtitle Translation of *My People, My Country*

As to the subtitle translation of the film *My People, My Country*, though few researchers instantly paid attention to the exploration of its subtitle translation because it was just launched in 2019, several researchers discerningly attempted to investigate the subtitle translation from various kinds of theories.

Liu (2020) explained how the four major techniques of translation variation were manifested in the subtitle translation of *My People, My Country*. Wang and Shen (2020) amplified the translation tactics of manifold story titles in the film based on the "three-dimensional transformation" principle of ecological translation. Tian (2021) dissected the application measures of Gideon Toury's translation norm theory in *My People, My Country*. In addition, Zhang (2020) first deployed multimodal discourse analysis theory to research *My People, My Country*. She worked on the study of the poster of The Eve, one of the story titles of the film, under the framework of visual grammatical analysis of multimodal discourse analysis. Later, Li explained how the three multimodalities of image, sound, and music have a bearing on the effects manifested in the film.

In a nutshell, all these foregoing research achievements accomplished by previous scholars or researchers have not only expanded the study territory of multimodal discourse analysis theoretically, laying a foundation for the subsequent explorations concerned, but also brought the theory to bear upon dynamic source materials practically, endowing real-life products with more ingenious analytical paths. Nevertheless, by combing through previous literature about both the theory and the artefact, the author noticed a handful of limitations that entail further improvement.

Theoretically, as a new academic sphere, multimodal discourse analysis is still situated in a rudimentary research phase. For instance, there are several primal deficiencies. Firstly, the studies are mainly centred around theoretical studies and there is insufficient empirical study. Secondly, domestic research is currently overfocusing on static texts such as posters, caricatures, etc., while there isn't enough research dynamically on foreign language teaching as well as subtitle translation with respect to multimodal discourse analysis. Thirdly, the definition of the relationships among different modalities requires more deep and scientific discussion.

Practically, for all applied research with regard to subtitle translation of *My People, My Country*, in some senses, there are still a few deficiencies in this research. For example, for some reasons, only a slight fraction of translation theories have been deployed. As mentioned above, these theories encompass perspectives on ecological translation, translation norms, and translation variation as well. Apart from the translation theories touched on above, only the linguistic theory—multimodal discourse analysis has so far been put into analytical research of the subtitle translation of *My People, My Country*.

Therefore, in order to achieve more substantial and significant successes in the realm of multimodal discourse analysis and subtitle translation of *My People, My Country*, both creative studies and practical explorations are indispensably required.

3. Theoretical Framework

3.1 Introduction of Multimodal Discourse Analysis

Before furthering the discussion on multimodal discourse analysis, relevant notions require elaboration for the sake of deep understanding. As to mode, Professor Zhu (2007, pp, 82-86) from Xi'an Jiaotong University considered it as the channel or intermediary of communication, which includes symbolic systems like languages, techniques, images, colours, and music; Professor Gu (2007, pp. 3-12) defined it as "the way people interact with their external environment (e.g. people, machines, objects, animals, etc.) through their senses (e.g. sight, hearing, etc." Common modes are composed of visual, auditory, tactile, smelling, and tasting modalities. Regarding multimode, it is elucidated by Kress & Van Leeuwen (1996) to be "a meaning combination of different semiotic systems". In addition, Gu (2007, pp. 3-12) assumed that mono-mode occurs when only one sense is utilised in interactions with the outer environment, two modalities when two senses are used, and multimodalities when three or more senses are involved. In practice, with various modalities deployed in communication, the audience can receive information through a few senses and get that information more successfully.

Currently, the most felicitous theoretical pattern for multimodal discourse analysis exploration is based upon the theory of systemic functional linguistics (SFL), as SFL confers upon researchers a viable theoretical framework for carrying out studies on multimodal discourse analysis.

Systemic functional linguist Halliday (1978) once described "language as a social semiotic". According to this view, he started to observe activities that transmit the meanings of mankind and study how these meanings are spawned in human life. Also, he illustrated three meta-functions of SFL—ideological, interpersonal, and textual. By virtue of this approach, Kress and Van Leeuwen then extended their research perspectives to the meaning dissemination of visual modes, which aroused academic interest in multimodal discourse analysis. Within the perspective of multimodality, linguistic symbols have been gradually put into a larger semiotic framework. This has been vindicated by Jewitt (2009), who asserted that "linguistic symbols are just one modality of the multimodal network." Thus, as resources to bring forth diverse meanings, non-verbal semiotic modalities, including images, sounds, colours, and others, are put in the equivalent position to verbal symbols.

Based upon systemic functional linguistics, Zhang Delu, a prominent scholar at home, brought forward a new framework of multimodal discourse analysis. This framework comprises four dimensions: cultural level, contextual level, content level, and expression level (Zhang Delu, 2009, p. 24-30). As can be seen in the figure below, each level includes a lot of subordinate concepts.
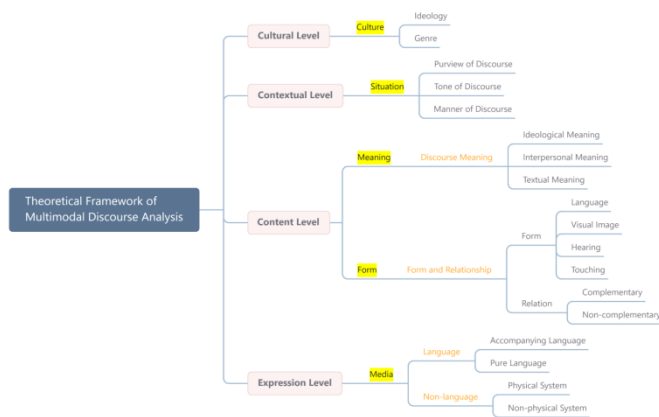
Figure 1: Theoretical framework of multimodal discourse analysis (Zhang Delu, 2009)

3.2 Cultural Level

According to Zhang Delu's (2009) analysis, cultural level is extremely important in enabling multimodal communication. To put it another way, the cultural level plays a pivotal role in the course of the ongoing transmission of multimodal ideas. Aside from that, the cultural level influences the choice of communicative techniques, forms, and traditions. Situational context, which belongs to the contextual level, will be deprived of persuasiveness and forceful strength. As for the level of culture, it constitutes ideology, which consists of human thinking modes, personal customs of life, philosophy of living, and all the hidden rules in society, and communicative programs or structural potentials—genre, which can be deployed to actualize this kind of ideology in a specific way (Zhang Delu, 2018, p. 731–743).

When it comes to the definition of language, linguist and semiotician Saussure once defined language as "a semiotic system that conveys notions" (Saussure, 1916). Chen Yuan (1984) regarded that as a kind of social phenomenon, language functions as the most significant instrument for mankind and is the direct realization of people's thoughts. More than a social phenomenon, language also functions as a mirror of certain cultural phenomena. Language is a manifestation of various cultural diversifications. Imbued with the thoughts and cultures of a nation, language is gradually integrated as one of the components of a special culture. At the same time, it can diffuse this national culture through culture itself (Chen Chen, 2015, p. 155+148). In conclusion, language is the carrier and expressive form of culture and cultural backstories.

To some extent, however, on account of discrepancies in people's thinking patterns, history, view of morality, religion, ideology, etc., differences in languages are inexorable. Therefore, with international interaction becoming more and more frequent and necessary, especially in the context of growing globalization, people from many countries find language to be a hindrance to understanding different cultures in the world. Consequently, this leads to tremendous inconvenience in cross-cultural communication and is detrimental to the process of multilateral cultural intercourse.

A relatively efficient way to address this problem is by translation. Translation involves the work to render a language into another. Catford (1965) believes that translation,

language and culture are interdependent, and translation is capable of both facilitating people's comprehension of the meaning from external culture and the communication of language.

With the advent of technology and multimedia, subtitle translation is currently widely employed as a method to bridge the gap between cultures. When translators are to achieve a satisfying effect in transmitting meanings and underlying implications, aside from complying with specific translation theory and policies, they should also notice that cultural level is at dominant play. Often, translators will make an explanation of the implicit cultural notions on the screen of films, televisions, and other audio-visual media in the form of subtitles. In practice, translators are required to have an all-around command of foreign audiences' demands for understanding the cultures of their nation and their capacities of receptivity (Yin Lin, 2008, p. 6-12). By doing so, audiences from other cultural backgrounds can make sense of what cultural meaning the screen is going to convey.

3.3 Contextual Level

In terms of context, it refers to the real situation when vocal behaviors take place, made up of factors including participants, time, location, formality extent of communication, topics, media, and participants' identities (Hou Guojin, 2003, p. 22-26). Besides, there are other concepts concerning context. For example, subjective context means the subjective factors of the communicator, such as personality, hobbies, emotions, etc., especially the specific state of mood. Objective existence in objective context communication such as specific spatial-temporal factors and various complex socio-cultural situations as the background. Then, more scholars in the domain decided to further their probes into research on context.

In terms of the relationship between culture level, context level and content level, the first level is demonstrated by the second while the second determines the last. From the functional linguistic angle of Zhang Delu (2009), communication is supposed to be under the restriction of contextual elements. These elements must be decided by the concrete purview of discourse, tones of discourse and manners of discourse. Accompanying this process of setting limits, the genre will be put into effect with a specific communication paradigm.

As one of the most important media for culture dissemination, the film is surely composed of auditory and visual components. More than that, the context, in which characters' interaction with others, action or articulation, deliberately-played background music and the scene purposefully set by the film director is applied to show some purposes or ideas, must also be taken into consideration. Sometimes, however, owing to nuanced distinctions of culture and social background between countries, context should also be interpreted by translators through the style of subtitle text. In this case, translators are requested to not only deal with culture level but also pay attention to context level. They have to realize the meaning of discourse in certain situations for the purpose to get the audience across the communication meanings of a particular context on the

screen.

3.4 Content Level

Generally speaking, the sub-framework of the content level mainly covers two tributary levels—meaning and form. The meaning level, which represents discourse meaning, is divided into three sections. They encompass ideational meanings, interpersonal meanings, and textual meanings, which are de facto derived from Halliday's (1978) meta-functions.

The form level is made up of forms and their relationships. Often, there are four forms, including language, visual images, hearing, and touching. They are all interconnected and manifest form features. Approaches to achieving meanings of modes vary differently. To be more specific, language form substantially sends its meaning to the audience through the lexical and grammatical systems; visual image through ideographic form and visual grammatical system; hearing through ideographic form and auditory grammatical system; and touching through ideographic form and feeling grammatical systems (Zhang Delu, 2009, p. 731–743).In general, the critical element in selecting multimodal discourse is concerned with how to achieve the desired effect when a cohort of modes collaborate with one another, and such inter-modal cooperation is embodied in form aspect and grammatical aspect in particular.

As for relations between these forms, complementary relations and non-complementary relations are the most appropriate pairs. This is how the complementary relationship can be translated—under no circumstances can a separate form or mode fully actualize its discourse meaning unless it coordinates with the rest of the auxiliary forms. Against this background, a single form accompanied by other complements gives birth to this complementary connection. Regarding the antithesis, it refers to the non-complementary relationship.

3.5 Expression Level

Professor Zhang (2009) contrived the construction of the expression level as he concluded it as the language and non-language forms within this framework. This level is in association with a category of media pattern employed by the speaker to convey the meaning of discourse. Being an acoustic correspondence, the media form is the material form in which discourse finally appears in the material world (Zhao Bing & Xiong Changqin, 2019, p. 97-102). Language media is composed of accompanying language and pure language. The former constitutes diverse media ranging from acoustic volumes to intonation, and from calligraphic typefaces to spatial layouts; the latter encompasses sound and characters. Multimodal discourse analysts hold that, between the two sorts of language, the accompanying one takes a fundamental position, which probably has a direct bearing on the conversion of the total meaning of discourse at times. Aside from typed words or written symbols, which refer to the language media of discourse, when people want to diffuse their own ideographic ideas, they have to make use of their own physical motions and other non-verbal devices, which are certainly the non-language media. Non-language media

comprises physical and non-physical systems. Under the structure of a non-language system, equipment and environment are kernel factors at play (Huang Huihui & Dujing, 2019, p. 133–135).

4. Case Study

4.1 Introduction to *My People, My Country*

The film *My People, My Country* tells China's main historical moments, national memories, and head-on collisions with the evolutionary new era. As well, it brings all of the audience back to the extraordinary years of the newly-established People's Republic of China in the past seven decades, making them deeply appreciate the connection of binding souls between the main characters and China and arousing the simplest patriotism in the audience's heart. *My People, My Country* is not only a tribute paid to the 70th anniversary of the establishment of our motherland, but also a romantic letter to it, with characters' stories, their experiences, and themselves confessing love to China.

In the form of collective creation, the film is directed by seven directors in seven episodes, which are filmed in different locations across the vast territory of China. The seven clips include: *The Eve, Passing By, The Champion, Going Home, Hello Beijing, The Guiding Star, and One For All.* A brief synopsis of them is as follows:

The Eve is set on October 1, 1949, the eve of the founding of the People's Republic of China. In order to ensure the smooth rise of the national flag at the founding ceremony, electric flagpole designer and installer Lin Zhiyuan (starred by Huang Bo) races against time during a thrilling sleepless night. Finally, he successfully secures the founding ceremony of the People's Republic of China. At the same time, the flag bearer, Comrade Fang (by Geng Le), manifold team members involved in the founding ceremony, and the masses work together, overcoming one problem after another, to guarantee the smooth flying of the five-star red flag over Tianmen Square. *Passing By is set in 1964, due* to the extremely high confidentiality of atomic bomb R&D, colleagues do not even know each other's names. Gao Yuan (by Zhang Yi), a researcher on the national defence science and technology front, has not contacted his family for 3 years. Between the love of the state and the love of couples, that generation of researchers choose the former, so there is only passing-by in their lives without reunion. *The Champion is set on* August 8, 1984, the Chinese women's volleyball team won the gold medal at the 1984 Los Angeles Olympic Games. On August 8, 1984, the Chinese women's volleyball team won the gold medal at the 1984 Los Angeles Olympic Games. At the same time, in Shanghai's Shikumen Alley, black and white televisions were set up within the aisle of the alley, and neighbors gathered together to watch the exciting game show. When the Chinese team scores, cheers seem to penetrate the sky. *Going Home* is set on July 1, 1997, when Hong Kong returned to the motherland. To make sure that the Chinese flag flutters over the sky in Hong Kong at the right time, flag bearer Zhu Tao (by Du Jiang) practices hard without negligence, while female police sister Lian (by Hui Yinghong) conscientiously patrols for local peace, diplomat An Wenbin (by Wang Luoyong) negotiates with the British over 16

seconds to punctually play the national anthem, and watch repairer Brother Hua (by Ren Dahua) exerts his exquisite skills to ensure the Hong Kong police's watch runs precisely. In *Hello Beijing*, taxi driver Zhang Beijing (by Ge You) accidentally gets a ticket for the opening ceremony of the Beijing Olympic Games. Amid showing off his ticket, Zhang comes across a series of hilarious experiences. In The Guiding Star, since vagrants Wo Dele (by Liu Hao Ran) and his younger brother Ha Zhabu (by Chen Feiyu) meet former Banner Manager——Director Li (by Tian Zhuangzhuang), they have undergone drastic changes. Under the guidance of Li, they witness the successful landing of the Shenzhou 11 spacecraft, and also see the daytime meteor in the fable of the grassland, which greatly impacts and baptises their hearts. In *One For All*, Lv Xiaoran (by Song Jia), the best female pilot in the PLA Air Force, is unexpectedly ordered to be the replacement pilot for the military parade commemorating the 70th anniversary of the end of World War II, which is held on September 3, 2015. When a fighter aircraft takes off, Lv Xiao Ran witnesses this glorious moment and also fulfills her mission.

Because of the colloquial and distinctive language characters, which are humorous and popularized in *My People, My Country*, film translation work personnel should coordinate the relationship of subtitle translation with other multi-modes and apply precise translation tactics according to characters' language features from the perspective of multi-modal discourse analysis. In this way, they can better help foreign audiences understand and appreciate the cultural landscape and strong patriotism that are depicted in the seven stories. In this chapter, the film will be analyzed according to the four levels of the theoretical framework of multi-modal discourse analysis.

4.2 Cultural Level

Successful subtitle translation requires not only translators to follow relevant translation principles and adopt appropriate translation methods, but also to deal with cultural issues in the process of translation (Jin Fang, 2007, p. 71).

As an intuitive means of cultural transmission, film carries copious cultural connotations. Subtitle translation is not only the conversion of language, but also the transmission of cultural information because "language bears cultural genes" (Tian Chuanmao & Wang Feng, 2017). As a result, as the translator embarks on rendering the source language into the target languages, he or she ought to be equipped with cultural consciousness and realize that subtitle translation is a distinct course across linguistic and cultural exchanges; then, the translator should try to get over impediments caused by cultural differences, so as to ensure the smooth realization of the emotional communication.

In addition, in the process of translation, sometimes for the purpose of delivering outstanding cultural characters to the target audience, translators are suggested to take advantage of specific scenes on the screen in order to make some cultural intentions clearer to the listeners, so as to achieve echoes at the cultural level (Xu Xiaoyan, 2020, p. 123).

However, cultural differences between different language countries exist objectively, which makes specific cultural images or connotations difficult to be understood by an audience with different cultural backgrounds. Under the call of the Chinese Culture Goes Out Strategy, Chinese film subtitle translators should find similarities in different cultures, consolidate efforts to make the cultural connotation of the original text much clearer and more transparent, and make the target audience understand the cultural images and connotations conveyed in the subtitle as much as possible in the appropriate time. The instantaneity of film and television works determines that subtitle translation must be short and clear. If the translated subtitle is problematic to make sense of, the fluency of watching the film will be greatly sabotaged, and the audience's grasp of the integrity of the movie plot will also be affected.

Example 1 (00:13:17):

The Source Version: 他巴不得我们搞不好呢
The Target Version: he'd love to see us blow it.

The plot concerning example 1 takes place in the quarrel between Lin Zhiyuan and Lao Fang in the first clip *The Eve*. The night before the founding ceremony, flagpole designer Lin Zhiyuan worked with his assistant Lao Fang to prepare the electric flag-raising device. Due to a shortage of manpower, his assistant Lao Fang suggested Lin Zhiyuan to seek the help of the team group. As the head of the manual team was in competition with Lin Zhiyuan, Lin thought that the manual team would not offer them help, so Lin said this word: "他巴不得我们搞不好呢", which according to literal translation can be rendered into "he'd rather we didn't get it right." However, in the subtitle translation, it has been translated into "he'd love to see us blow it." Indeed, "blow it" is a slang which means "to ruin, mishandle, or fail to capitalize on an opportunity." By using this slang, the audience can sense the unsupportive attitude of the head of labor team towards Lin Zhiyuan's work, and even the idea of gloating, as well as Lin's non-expectations of the help from the labor team. Combined with the given image and the characters' actions, foreign audiences are very likely to know the underlying meaning of "他巴不得我们搞不好呢" and understand what the characters were doing.

Example 2 (00:13:33):

The Source Version: 你个穷酸秀才
The Target Version: You pathetic egghead,

Example 2 is from the same clip *The Eve*. After Lin said "他巴不得我们搞不好呢"and complained that the labor team leader is so narrow-minded, the labor team retorted that "你个穷酸秀才", which can be literally translated into "a poor and pedantic scholar". The subtitle translator rendered it into "You pathetic egghead." The slang "egghead" is the meaning of "a highly academic or studious person." Compared with the literal translation, this subtitle translation can be more familiarized by the audience and be in line with the target culture background. Subtitle translation at the cultural level largely helps bridge the cultural abyss among people from different countries and conveys meanings to the audience.

4.3 Contextual Level

In discourse, especially in film and television, languages, contexts, images, and other factors are very important to deliver meanings so as to achieve proper interaction with the audience. Zhang Delu (2009) once pointed out that a relationship between mode and context can be regarded as the relationship between positive and negative modes. Actually, the film is composed of a series of dynamic pictures, and each of them is logically linked to the previous one, which determines the strong dependence of the film on the context, a sort of feature of multi-modality. According to Hu Zhuanglin (2007), context can be categorized into text context, situational context, and cultural context, among which situational context is an essential factor affecting subtitle translation. Translators can make full use of contextual information to appropriately complement or excise text information so that the discourse meaning on the screen is presented naturally and smoothly.

Example 3 (00:24:48):

The Source Language: 天桥的手艺 擎好吧您内
The Target Language: I know what I'm doing, you have nothing to worry about.

Example 3 is from the first section *The Eve*. When Lin Zhiyuan was allowed to enter Tianmen Square, he and his security officer Du rushed into the place without the instant permission of the military guardians. Then, he asked a young man to climb up along the flagpole in order to fasten the device on the top of the pole. Lin, together with Lao Du and other workers, fixed the flagpole with rope and tightened it under the flagpole to keep it stable and ensure the safety of the young person on the pole. When the man nearly climbed down to the pedestal of the flagpole, Du asked whether he had fastened the device. At that time, the young man replied: "天桥的手艺, 擎好吧您内." Actually, the words "天桥的手艺" refers to the skills or expertise of a certain group of artists often seen on foot bridges. In history, this cohort of people were named bridge artists, a sort of artistic community in the long term showing up among the bridge areas in Beijing, the birthplace of many folk arts. These artists have always been regarded as excellent in their outstanding skills. "擎好吧您内" is characterized in the dialects of Beijing, producing a sense of gentility. Judging by the previous contexts of the film, nothing about bridge artists was mentioned. The young man's intention was to show his confidence in his skills and reassure Lin Zhiyuan and other workers.

As well, the founding ceremony was imminent, foiling a quite tense atmosphere when the spectators were trying to make sense of the skills of the bridge artists. Thus, the subtitle translation of "I know what I'm doing, you have nothing to worry about." is more connected to the original context, enabling the audience to more easily understand the core meaning of the discourse.

4.4 Content Level

Content level is divided into discourse meaning level and form level. The level of discourse meaning includes the concept of ideational meanings, interpersonal meanings, and textual meanings. The formal level includes the formal features of different modes such as language, sense of images,

sound, and sense and their relations with each other. At the level of discourse meaning, as subtitle translation is restricted by time and space factors, translators usually deploy the reduction method to delete interpersonal meaning and textual meaning, but less ideational meaning (Yang Wenwen, 2012, p. 155-156). It can be noticed that in the subtle translation of *My People, My Country*, modal particles are usually omitted, because the deletion of these words does not affect the transmission of the overall meaning of the film discourse, and the audience can still understand the meaning of the original text by combining the film pictures or background sound effects.

Example 4 (1:04:18):

The Source Language: 嘘 听不清啊
The Target Language: It's hard to hear.

Example 4 appears in *The Champion*. When residents living in Shanghai Shikumen came together to watch the exciting women's volleyball tournament, there was an uproar of noise from crows. On this occasion, the spectators in front of the only television could hardly hear the commentary of the volleyball competition. In spite of the omission of the modal particle "嘘", the audience can still understand how eager the residents were to listen to the commentary and hope that our Chinese women's volleyball team would win the final games. In the movie, the characters were very excited. The background music also highlights the mood of the people, and everyone's expressions were very eager. The subtitle translation properly coordinates with these different modes, obviously making the audience know how thrilling the game was and how excited the residents were.

Example 5 (1:04:33):

The Source Language: 哎呀 你看什么地方呢
The Target Language: What are you looking at?

Example 5 also exists in *The Champion* taking place after Example 4. Xiao Dong was then on the proof to fix the antenna of television. Several minutes later, Xiao Mei, the girl who Xiao Dong liked, was to take the taxi and move to the next place. It was difficult to get downstairs immediately, so he asked Xiao Mei to wait for him just for a while. When Xiao Mei looked at the proof, her mother felt there was not enough time to wait and they might miss the set time. Similar to the translation of Example 4, the modal particle "哎呀" was also excised in the subtitle translation. At that moment, everyone was very nervous—the residents were nervous about the game situation of the women's volleyball team, and Xiao Mei's mom was nervous about the time to catch the taxi. Their voice, their expression, the background music, along with the omitted subtitle translation, exactly delivered the discourse meaning to the listeners together.

4.5 Expression Level

The level of expression includes verbal media and non-verbal media (language system and non-language system). Language media generally refers to sound symbols and characters, and media forms accompanying language (such as volume, tone, timbre, font shape, size, and layout) play an auxiliary,

complementary, or reinforcing role in conveying semantics. In modern society, audiences prefer to watch movies in their original sound with subtitles in the translated language. This requires that the subtitle content of the target language be concise and the time required to complete the information transfer be roughly the same as that of the source language, and there needs to be enough time for the audience to read the information.

Example 6 (01:21:33):

The Source Language: 展示文明之师 威武之师的风采
The Target Language: Let's show the world our honor and our strength

Example 6 is from the clip *Going Home*, what the military training officer said to encourage the ceremonial brigade before the handover ceremony of Hong Kong's return. The subtitle translator rendered "展示文明之师 威武之师的风采" into "Let's show the world our honor and our strength" in which the translation of "文明之师" and "威武之师" were omitted. It is reasonable from the perspective of the duration of the film and the collocation of the text content, its purpose is to remove redundant information, appropriately set the length of subtitles, reduce the burden of watching the film, and help the audience understand the content of the section without the meaning of the film's discourse undermined. Besides, the subtitles on the screen are aligned with the audio-visual image. This is a small task for the media system, but it does help to improve the quality of the subtitle translation and makes it easier for viewers to enjoy the film.

5. Conclusion

The thesis tries to probe subtitle translation from the perspective of multimodal discourse analysis. Through applied analyses, the paper explores the merits of multimodal discourse analysis on subtitle translation. With these efforts, it has been discovered that on some occasions, hardly any monotonous modality is capable of thoroughly disseminating the whole meaning of the discourse on the screen, and the meaning can only be conveyed properly and effectively with the combination of different modalities. Copious modalities are mainly composed of sound, images, tones, gestures, volumes, touching, and so on, which contribute to a more accurate expression of the meaning. In addition, as to which modality is to be selected, it must take the intention of communication into account.

As previously stated, the integrated framework of multimodal discourse analysis extends beyond the boundaries of traditional cultural and linguistic perspectives. By largely deploying modalities like sounds, verbal factors, and images, multimodal discourse has achieved a leapfrog in research on subtitle translation. In this case, more and more spectators find it easier to figure out the meaning of media works such as films and advertisements. Therefore, multimodal discourse analysis is viable in subtitle translation by applying the theory of multimodal discourse analysis. Notwithstanding the progress in investigating the subtitle of *My People, My Country* from the perspective of multimodal discourse analysis, limitations reside in this thesis. For instance, even though this paper aims to have an all-round understanding of

multimodal discourse analysis in *My People, My Country*, on account of the shackle of time and space, there is less likelihood of paying attention to all the modalities of this film. What's more, owing to the limit of time, the thesis can scarcely present all the corresponding examples to fully demonstrate the multimodal discourse analysis on subtitle translation. In addition, this thesis only delves into the practical application of the theoretical framework of multimodal discourse analysis in subtitle translation from a macroscopic perspective, while failing to analyse the subtitle translation of *My People, My Country* from a more concrete view in this theoretical framework.

In spite of some problems remaining to be addressed, the thesis still achieves strides in both applied and theoretical study. The conclusion of this paper is that multi-modal interaction can more effectively convey the meaning of discourse in text and audio-visual media. Thus, multimodality can be widely utilised by enormous numbers of individuals and units in the process of teaching students in the classroom or online, disseminating films to foreign countries, and promoting the profile of companies to their target customers and partners. In the meantime, in the course of rendering the source language into the target language, the subtitle translator needs to focus on cultural backgrounds and contexts, and take advantage of various modalities existing on the screen to refine the translation version. That is to say, the translator is to achieve an equilibrium between visual images, sounds and discourses. Only by taking all these components into account can the film entitatively achieve a satisfying communicative effect.

References

[1] Catford, J. (1965). *A linguistic theory of translation*. London: Oxford University Press.
[2] Chen, Y. (1984). *Sociolinguistics*. Hong Kong: The Commercial Press (HK) Ltd.
[3] Chen, C. (2015). A brief discussion on foreign language teaching and foreign culture education in universities. *New West*, (22), 155+148.
[4] Gu, Y. (2007). On multimedia learning and multimodal learning. *Technology Enhanced Foreign Language Education*, (2), 3-12.
[5] Harris, Z. S. (1952). Discourse analysis. *Language*, 28(1), 1-30.
[6] Halliday, M. A. K. (1978). Language as social semiotic: *The social interpretation of language and meaning*. Baltimore: University Park Press.
[7] Hou, G. (2003). Dynamic context and contextual negotiation. *Foreign Language Education*, (01), 22-26.
[8] Hu, Z. (2007). Multimodalization in social semiotics. *Language Teaching and Linguistic Studies*, (1), 1-10.
[9] Hu, Y. (2014). On the translation of multimodal text based on the theory of text typology. *Journal of Suzhou University*, 6(4), 52-55.
[10] Huang, H., & Du, J. (2019). The translation strategies of documentary subtitles from the perspective of multimodal discourse analysis: A case study of Amazing Henan. *Journal of Lanzhou Institute of Education*, 35(02), 133-135.
[11] Jin, F. (2007). Preliminary study on subtitle translation strategies. *US-China Foreign Language*, (05), 71.

[12] Jewitt, C. (2014). *An introduction to multimodality: The handbook of multimodal analysis*. London: Routledge.

[13] Bateman, J. A., & Wildfeuer, J. (2014). A multimodal discourse theory of visual narrative. *Journal of Pragmatics*, 74, 180-208.

[14] Kress, G., & Van Leeuwen, T. (1996). *Reading images: The grammar of visual design*. London: Routledge.

[15] Kress, G., & Van Leeuwen, T. (2001). *Multimodal discourse: The modes and media of contemporary communication*. London: Routledge.

[16] Li, Y. (2000). Strategies for translating subtitles. *Chinese Translators Journal*, (7), 38-40.

[17] Li, Z. (2003). Social semiotic approach to multimodal discourse. *Foreign Languages Research*, (5), 1-8.

[18] Li, J. (2020). A multimodal discourse analysis of dynamic discourse: A case study of the film *My People, My Country*. *English Square*, (22), 45-48.

[19] Liu, X. (2020). The application of translation variation theory in film subtitling: A case study of *My People, My Country*. *China Journal of Multimedia & Network Teaching (Upper Edition)*, (23), 247-248.

[20] Norris, S. (2004). *Analysing multimodal interaction: A methodological framework*. London: Routledge.

[21] O'Toole, M. (1994). *The language of displayed art*. New Jersey: FDU Press.

[22] Saussure, F. (1916). *Course in general linguistics*. London: Duckworth.

[23] Tian, C., & Wang, F. (2017). *Translation and culture*. Beijing: China Social Science Press.

[24] Tian, Y. (2021). A study on subtitling strategies from the perspective of Toury's norm theory: A case study of the film *My People, My Country*. *Culture Industry*, (8), 89-91.

[25] Wei, Q. (2009). On mode, medium, and modality in multimodal discourse. *Foreign Language Education*, (4), 54-57.

[26] Wang, Y., & Shen, M. (2020). An analysis of the English translation strategies for the story titles in the film *My People, My Country* from the perspective of eco-translatology. *English Square*, (10), 14-17.

[27] Xie, S. (2020). On English translation of My Motherland and Me from a perspective of literary stylistics. *Journal of Xiangyang Polytechnic*, 19(4), 131-133.

[28] Xu, X. (2020). A study of subtitle translation from the perspective of multimodal discourse: A case study of the film Hero. *Journal of Harbin University*, (03), 123.

[29] Yang, W. (2012). Subtitling translation of Gone with the Wind from a multimodal discourse analysis perspective. *Movie Literature*, (19), 155-156.

[30] Yin, L. (2008). The translation strategy of cultural terms in film subtitles: An analysis from the perspective of functional translation theory. *Research in Foreign Language and Literature*, 8(02), 6-12.

[31] Yu, X., & Geng, Q. (2018). Studies of visual texts translation: Theories, problematics, and methods. *Foreign Languages and Their Teaching*, (3), 77-87+145.

[32] Zhu, Y. (2007). Theory and methodology of multimodal discourse analysis. *Foreign Language Research*, (5), 82-86.

[33] Zhang, D. (2009). On a synthetic theoretical framework for multimodal discourse analysis. *Foreign Languages in China*, (1), 24-30.

[34] Zhang, D. (2018). A systemic-functional synthetic framework for multimodal discourse analysis. *Modern Foreign Languages*, 41(06), 731-743.

[35] Zhao, B., & Xiong, C. (2019). A multimodal perspective analysis of the English version of the Jiangxi global promotion video. *Journal of Jiangxi University of Science and Technology*, 40(04), 97-102.

[36] Zhang, D., & Zhao, J. (2021). Does multimodal discourse analysis require the analysis of multimodal grammar? *Contemporary Rhetoric*, (2), 26-36.

[37] Zhang, M. (2020). A multimodal discourse analysis of film posters from the perspective of visual grammar: A case study of the poster for the film *My People, My Country* (The Eve). *Dazhong Wenyi*, (16), 71-72.