

# Subtitle Translation from the Perspective of Multimodal Discourse

-- Take the Movie *Harry Potter and the Philosopher's Stone* as an Example

Peiyang Lin<sup>1</sup>, Dijiang Luo<sup>1,\*</sup>

<sup>1</sup> School of Foreign Studies, Guangxi University of Science and Technology, Liuzhou, China

\* Corresponding author: Dijiang Luo (Email: dijiangluo@163.com)

---

**Abstract:** Film and television works are multimodal discourse composed of a variety of symbol systems such as text, sound and image, so the audience's various senses can be mobilized at the same time when watching movies. Starting from the perspective of multimodal discourse analysis, this paper applies Delu Zhang's theoretical framework of multimodal discourse analysis to analyze the subtitle translation of *Harry Potter and the Philosopher's Stone* from four aspects: culture, context, content and expression, so as to understand the influence of multimodal on subtitle translation.

**Keywords:** Audio-Visual Translation, Multimodality, Subtitle Translation.

---

## 1. Introduction

Audio-visual translation has almost developed with the Chinese film history. With the advent of the Internet era in the 1990s, film and television works began to become accessible, and the upsurge of folk translation began to rise, which led to the birth of many well-known grassroots subtitle groups. As a bridge of cross-cultural communication, film and television translation undertakes the responsibility of crossing the barriers of different cultures and languages, and realizes the cultural transmission of film and television works in the world. There are two basic forms of audio-visual translation: dubbing and subtitle translation. At present, subtitle translation from network is still the most popular and widely spread subtitle translation. The main force includes the old classic subtitle groups such as Renren subtitle group, Ragged Bear subtitle group, Holy City Homeland subtitle group and other excellent subtitle groups. At the same time, the streaming media platforms represented by Youku and Sohu also made indelible contributions to subtitle translation. They introduced a large number of original Korean and American TV series, which became the main front for Korean and American dramas to spread in China at that time; And because of the high quality of subtitle translation, many translated versions by these platforms are still classics. In general, in the decades of development of film and television translation, although China still lags behind the West in the academic research of audio-visual translation, our translation practice can keep pace with the development of the times. Therefore, this is also a desirable aspect of the development of Chinese audio-visual translation. Different from traditional literary translation and drama translation, film and television works are not only pure texts, but also contain color, music, performance, etc., which make the multimodal characteristics more obvious and constitute the multimodal discourse of subtitle translation. Therefore, studying subtitle translation from the perspective of multimodality could evaluate the quality of translation in a more comprehensive way.

## 2. Multimodal Discourse Analysis

Barthes was the first abroad to study the multimodal discourse analysis. In *Rhétorique de l'image*, he studied the interaction between images and language in expressive meaning; Kress & Van Leeuwen (1996; 2001; 2003) studied the relationship between multimodality and media, and specifically discussed the phenomenon of multimodal phenomena which regularly expresses meaning, including visual images, color, newspaper layout design and the role of different media; O'Halloran (2004) not only studied the theoretical construction of multimodality, but also studied multimodality in mathematical discourse.

The research on multimodality in China started relatively late. It was not until 2010 that multimodal discourse analysis received more attention from scholars. Zhanzi Li, Delu Zhang, Zhuanglin Hu, Yongsheng Zhu and other well-known domestic experts participated in the research on multimodal discourse analysis. Li (2003) first mentioned multimodality in *Social Semiotic Analysis of Multimodal Discourse*, and studied the theory of multimodality discourse with the theory of systemic functional linguistics; Hu (2007) studied multimodal symbols from the perspective of theory and practice, and distinguished multimodal semiotics from multimedia semiotics; Zhu (2007) studied the theoretical basis and research methods of multimodal discourse; Zhang (2009) proposed the definition of multimodal discourse on the basis of previous studies, and summarized the theoretical framework of multimodal discourse analysis.

Multimodal discourse refers to the phenomenon of using a variety of senses, such as vision, hearing, smell, taste, etc., to communicate through text, pictures, sounds, animation, video and other social symbols, which have their own systematic grammar, and can express meaning independently or construct meaning cooperatively. Multimodal discourse analysis brings nonverbal signs into the process of meaning communication and conducts research, including images, actions, eyes, gestures, etc., and also focuses on the relationship between linguistic signs and nonverbal signs. Zhang (2009) pointed out that the most suitable theoretical

model for multimodal discourse analysis is the theory of systemic functional linguistics, and accordingly proposed the four levels in the theoretical framework of multimodal discourse analysis: cultural level, context level, content level, and expression level. On the cultural level, the tradition, form and technology of communication are determined by the cultural level. As a mass media, film and television works are conscious cross-cultural communication activities filled with cultural meanings. Whether cultural elements are conveyed in film and television works to the audience or not directly affects the audience's acceptance and understanding of the film. Therefore, it is necessary to attach importance to the transmission of cultural elements when adopting translation strategies. On the context level, communication is restricted by the context factors which determined by the scope, tone and mode of discourse in the specific context. The content level includes the meaning level and the form level. On the meaning level, the translation is required to convey the information accurately. On the form level, the translator needs to take language, pictures, music and other factors into consideration to translate the original text in a succinct and understandable way. The expression level is the final manifestation of discourse.

### 3. Multimodal Discourse Analysis of Subtitle Translation of *Harry Potter and the Philosopher's Stone*

For subtitle translation, foreign scholars define it as translating the language information in the film and television media into another language, and simultaneously expressing the meaning of the source language with one or two lines of translated text on the screen (Gottlieb, 2001: 85-110). In the process of subtitle translation, considering the cultural differences, the translated content should adapt to the cultural traditions and norms of the target language. In the study of translation processing strategies, foreign scholars have also put forward some universal methods. Gottlieb (2009: 24) proposed that translation strategies should be adjusted according to different types of film and television works. Due to the special characteristic of film and television works, audio-visual translation is different from literary translation and drama translation. It brings the audience the real-time "reading experience" in both visual and auditory dimensions. There are many elements in film and television works that can become the object of multi-modal research, such as language, light, sound, color, dynamic pictures, etc., which influence each other and constitute the overall meaning of the film; In order to make audience understand the film, the joint effect of multimodal expressions is needed. Therefore, studying subtitle translation from the perspective of multimodal discourse analysis can help us break away from the traditional study on pure text, conducting a more comprehensive evaluation of advantages and disadvantages of subtitle translation; and propose more appropriate translation strategies based on the analysis of subtitle translation.

#### 3.1. Subtitle Translation on the Cultural Level

Zhang (2009) proposed that the cultural level is the key level to make multimodal communication possible. The tradition, form and technology of communication are all determined by this level. Without cultural level, situational context will not be explainable. Film and television work themselves are cross-cultural communication activities, so we

should pay attention to the transmission of cultural elements in film and television translation. For example, the original work of the film *Harry Potter and the Philosopher's Stone* is a children's literature that has achieved great success all around the world. Therefore, in the subtitle translation and production, the younger audience should also be taken into consideration. In the translation, both the translation principles and cultural phenomena should be followed.

(1) Harry: "Your father left this in my possession before he died. It is time it was returned to you. Use it well."

哈利: "令尊过世前把此物托付给我, 今物归原主, 请妥善使用"

("Your Highness entrusted this to me before his death. Now it is back to its owner. Please use it properly.")

In this scene, Harry received a Christmas gift from Dumbledore, which was the invisibility cloak he borrowed from Harry's father, and he wrote the above sentence on the card. Since Harry's father has died in battle, the formal translation of the letter can reflect Dumbledore's respect for Harry's father. "令尊" is a kind of honorific title for other people's father in the Chinese context. Therefore, the translation of "your father" into "令尊" is also in line with the Chinese habit of formal written language, so that the Chinese audience can directly feel Dumbledore's attitude towards Harry's father from the formal language.

#### 3.2. Subtitle Translation on the Context Level

Zhang (2009) proposed that, in a specific context, communication should be restricted by contextual factors, including contextual factors determined by the scope, tone, and the mode of discourse. Therefore, context will influence and restrict the semantic meaning, words, structural form and language style of language, that is, what people say under what circumstances. Since film and television works belong to audio-visual art, subtitle translation is more likely to be constrained by contextual factors, such as scene atmosphere, pictures, sound effects, clothing, tone of voice, facial expressions, body language and so on. For example:

(2) Harry: You knew? You knew all along and you never told me?

哈利: 你们知道? 怎么从来没告诉过我?

(You knew? Why didn't you ever tell me that?)

Petunia: Of course we knew How could you not be? My perfect sister being who she was.

佩妮: 我们当然知道。你也是那种货色, 就像我姐姐一个样子

(Of course we do. You're one of those people, just like my sister.)

In the film, Hagrid, the key keeper of Hogwarts, finally sends the Hogwarts admission letter to Harry on the sea island, thus revealing the true identity of Harry's parents as wizards and the real reason for their death. Harry was very angry when he learned that, and asked his aunts and uncles why they hid the truth from him. In this scene, Aunt Petunia mentioned that her sister, Lily, Harry's birth mother, also used to receive an admission letter from Hogwarts. However, according to the original text of the novel, only when she had the talent to learn magic could she receive an admission letter. In the family, only Lily received the letter, while Aunt Petunia did not. Therefore, Aunt Petunia's feelings for Lily have been very complicated for many years. While calling her a "freak" and resisting magic (this behavior also explains why Uncle Vernon destroyed Harry's admission letters before, while

Aunt Petunia did not mention Lily's identity or the cause of death), Aunt Petunia's feeling towards Lily was also mixed with the same desire to be the one chosen to be a wizard. When Aunt Petunia was talking, we could see from her facial expression that her eyes were full of disdain for Lily, but at the same time there were also tears in her eyes. Therefore, when translating "How could you not be", the subtitle group translated it as "货色", which is derogatory and accurately conveyed Aunt Petunia's emotions in the current context. If translate it literally into "How can you not be that kind of person", it would fail to accentuate Aunt Petunia's tone of disgust, and would make their relationship even stranger. The subtitle translation here is consistent with the context and the overall meaning of various modalities, which helps us understand the characters and the meaning of the whole film.

In addition, in film and television works, the picture background and language are also important information for the development of the story. Therefore, while translating, text information should be added or reduced according to the context, so that the audience can grasp the overall story. For example, when Hagrid took Harry to the Gringotts Bank to withdraw money, he handed a letter from Dumbledore to the chief goblin. The envelope was stamped with a red seal saying "Top Secret". Here, the camera gives a close-up of the letter. This is not an ordinary letter. In the following plots, it can be seen that the vault where Dumbledore entrusted to take things is of a very high security level, and it is the Philosopher's Stone that Hagrid was entrusted by Dumbledore to take away. Therefore, the "Top Secret" is the key information of the film, which is also the foreshadowing to promote the following plot. So the subtitles read the context information and translated the "最高机密", which means "top secret", and conveys the meaning of the film to the audience through the text mode helping the audience understand the development of the plot.

### 3.3. Subtitle Translation on the Content Level

The content level includes the meaning level and the form level. Zhang (2009) proposed that the meaning level includes conceptual meaning, interpersonal meaning and textual meaning restricted by scope, tone and mode of the discourse. Conceptual function includes experiential function and logical function, both of which are based on the speaker's experience of the external world and the internal world. One of the functions of language is interpersonal function, which reflects the relationship between people, or expresses one's own judgment and estimation of the possibility and frequency of things, or reflects the social status and closeness between speakers and listeners. The part of conceptual function and interpersonal function can be realized only when the speaker organizes them into discourse, then it forms the discourse function (Hu, 2005: 12). On the form level, the formal characteristics of different modes are interrelated and jointly reflect the meaning of discourse. "The form level in film discourse is mainly represented by music in auditory mode, language and other sound effects, and views and texts in visual mode." (Lv, 2012:36) Therefore, on the level of meaning, the subtitle translation should accurately convey the information; On the form level, the translation should consider the comprehensive effect of visual and auditory, such as the picture, music tone, color and other factors, and the translation should be as simple and easy to understand as possible.

A typical example of the meaning level is Professor Snape's debut in the first potions class for freshmen:

(3) Snape: There will be no foolish wand-waving or silly incantations in this class.

斯内普: 在我的课上, 不准乱挥魔杖, 不准乱念咒语  
(No waving wands or chanting spells in my lessons.)

In this scene, Snape suddenly opened the door, walked into the classroom quickly and said the above words. His face stern and his eyes sharp, his black robes flowing behind him in a crushing manner. The subtitle translation took into account the characteristics of many short sentences in Chinese, and combined with the characteristics of the characters and the tone of voice, the whole sentence of the original text was divided into three short sentences. In addition, "in this class" was not translated as "在课堂上" (in this class) but "在我的课上" (in my class). The imperative sentence pattern was used to express prohibition, which not only directly demonstrated the relationship between Snape and the students, but also established Snape's image as a grim authority, directly showing that he is not a teacher who would socialize with students.

(4) Hagrid: But one thing's certain. Something about you stumped him that night.

海格: 但有一件事很肯定。那天晚上你一定让他元气大伤

(But one thing's for sure. You must have made him lose so much power that night.)

In the film, Harry found that the wand suitable for him had the same core as Voldemort's wand, and their feathers came from the same phoenix. Harry was confused and ask Hagrid about his connection with Voldemort. Hagrid roughly told the story of the night when Harry survived, and mentioned that Voldemort was still missing. In the dictionary, "stump" means to baffle someone. Considering the plot of the film, we could know that on the night of Harry's survival, Voldemort deadly curse on Harry accidentally failed, and the spell bounced back to Voldemort himself, leaving his body in pieces and his soul nowhere to be placed. Since then, he has disappeared. No one has ever been able to survive under Voldemort's deadly curse, so no one knows exactly why Harry has survived. Therefore, it would be more accurate and more concise to translate "stump" into "元气大伤" in combination with the plot.

The form level in film and television works is mainly represented by music and language in auditory modality, as well as views and texts in other sound effects and visual angle modalities. For example:

(5) Voldemort: Why suffer a horrific death when you can join me and live?

伏地魔: 何苦自寻死路, 不如共享永生?  
(Why die, rather than live forever together?)

At the climax of the film, Harry entered the basement and confronted Voldemort. Voldemort threatened Harry to hand over the Philosopher's Stone that can make people immortal, but Harry refuses decisively. Voldemort made Professor Quirrell ignite the fire, cutting off Harry's only escape and beginning to surround Harry. In this scene, the music turns from calm to tense as Voldemort, the much-feared "Dark Lord," still wielded immense power, assuming a commanding posture and speaking with a fierce face, despite inhabiting Quirrell's body. The meaning of the lines is "Why do you have to suffer a terrible death since you can join hands with me to achieve eternal life". However, instead of literal translation, the translation directly adopts the antithesis translation instead of the logic of the original sentence, which directly reveals the unique temperament of Voldemort as a villain. The

subtitle translation combines the overall meaning constructed by visual and auditory meaning, and the image Voldemort was instantly established as a villain.

### 3.4. Subtitle Translation on the Expression Level

Zhang (2009) pointed out that, according to traditional linguistics, the media for meaning transmission mainly included sound symbols transmitted by sound waves and written symbols generated by pens, which were the main media for language to transmit information. Audio-visual subtitles are written symbols on the expression level, so the corresponding translation also appears in the form of characters. However, in the translation, the characteristics of language media, such as the voice and tone, the size and shape of subtitles, and the space layout of subtitles on the screen, should all be considered. Therefore, subtitle translation must complete the transmission of information in the same time as the source language does, and the transmission must be accurate. The subtitles of this film produced by Renren's subtitle group have basically designed to have the same length in Chinese and English. For example, when Harry asked Dumbledore how he was able to take out the Philosopher's Stone from his pocket, Dumbledore replied:

(6) Dumbledore: You see, only a person who wanted to find the Stone, find it, but not use it, would be able to get it.

邓布利多: 你看, 必须是想要找到魔法石, 找到它, 但不利用它的人, 才能得到它

(You see, it must be someone who wants to find the Stone, find it, but not use it, to get it.)

The subtitle production of this paragraph is particularly good at information segmentation. The key information is conveyed accurately, and at the same time, the length of Chinese and English subtitles is approximately the same, making the subtitles beautiful and orderly. The original text is a long sentence with long clauses and lots of parentheses. If it is translated as “你看, 只有想要找到魔法石、找到了又不利用它的人才能得到它” (You see, only those who want to find the Philosopher's Stone can get it if they find it and do not use it), the length of the subtitle cannot be roughly equal to that of the source language, and it will be difficult for the audience to spend more time to read the subtitles. Therefore, while ensuring the accuracy of the translation, the Chinese subtitle needs to control the length of the sentence to be about the same as that of the original text, facilitating the audience to access subtitles without causing confusion.

## 4. Conclusion

China has made great achievements in the practice of subtitle translation, but the theoretical research still lags behind the world's advanced level. From the above analysis

of subtitle translation, it can be concluded that multimodal discourse analysis has helped us establish a new way to deal with the relationship between language symbols and other symbol systems in the study of audio-visual translation, so that we can no longer just focus on the pure text when studying the translation, thus obtaining better translation strategies and methods. The fact that subtitle translation can be analyzed from the perspective of multimodality shows that subtitle translation is distinctive multidisciplinary. Therefore, audio-visual translation should receive more attention and be studied more extensively and deeply in the future.

## References

- [1] Kress, G. and Van Leeuwen, T. *Reading Images: The Grammar of Visual Design* [C]. London: Routledge, 1996.
- [2] O'Halloran, Kay L. *Multimodal Discourse Analysis: Systemic-Functional Perspectives* [C]. London: Continuum, 2004.
- [3] Xu Yihang. The Challenges of Film Translation from Multimodal Perspective [D]. Shanghai International Studies University, 2021.
- [4] Zhao Jirong. Subtitle Translation of British drama *Sherlock* into Chinese: A Multi-modal Discourse Analysis Perspective Based on Audio-Visual Translation [J]. *Journal of Lanzhou Institute of Technology*, 2017 (6): 127-132.
- [5] Chen Yuping, Zhang Caihua. Chinese Translation of English Film Subtitles: A Multimodal Analysis of the Relationship between Text and Text [J]. *Chinese Journal of Translation*, 2017 (5):105-110.
- [6] Sun Man, Yuan Jian. The Influence of Multimodality of Film Text on Subtitle Translation [J]. *Journal of Hubei Institute of Technology*, 2015 (1): 77-79.
- [7] Kan Feng. Film Subtitle Translation from the Perspective of Multimodal Discourse Analysis [D]. Qiqihar University, 2014.
- [8] Lv Jie. Research on Western Audiovisual Translation from 2000 to 2012 [J]. *Shanghai Translation*, 2013(4): 68-72.
- [9] Li Zhanzi, Lu Danyun. Multimodal Semiotics: Theoretical Basis, Research Approaches and Development Prospects [J]. *Foreign Language Studies*, 2012 (2):1-8.
- [10] Chen Hongyan. Subtitle Translation in Multimodal Mode -- A Case Study of Subtitle Translation in *Gone with the Wind* [J]. *Film Review*, 2011(19): 71-73.
- [11] Zhang Delu. Research on the Comprehensive Theoretical Framework of multimodal Discourse Analysis [J]. *Foreign Languages in China*, 2009 (1):24-30.
- [12] Zhu Yongsheng. The Theoretical Basis and Research Methods of Multimodal Discourse Analysis [J]. *Foreign Languages Journal*, 2007 (5):82-86.
- [13] Hu Zhuanglin. Multimodality in the Study of Social Semiotics [J]. *Language Teaching and Research*, 2007 (1):1-10.