

Subtitle Translation in the Online Promotional Video *CHN* for the International Image of Chinese Civilization From the Perspective of Multimodal Discourse Analysis

PENG Ying

Wuhan University of Technology, Wuhan, China

Multimodal discourse analysis has advanced the shift in subtitle translation studies from linguistic level to modal synergy. National image promotional videos, as highly condensed multimodal texts, hold significant research value for their subtitle translation in cross-cultural communication. Drawing on Zhang's multimodal discourse analysis framework, this paper analyzes the dynamic interaction and meaning co-construction among visual, auditory, and textual modalities in the international promotional video *CHN* on Chinese civilization. The findings contribute to a deeper understanding of the mechanisms underlying multimodal subtitle translation and offer insights for subtitling practices in the context of promoting Chinese culture globally.

Keywords: multimodal discourse analysis, subtitle translation, online promotional video *CHN* for the international image of Chinese civilization

Introduction

Amid the evolving global communication landscape, audiovisual texts have gained increasing prominence in cross-cultural exchanges. Documentary films, promotional videos, and other audiovisual discourses construct meaning through the synergistic interplay of multiple modalities like language, imagery, sound, and editing, making subtitle translation a process of cross-cultural recontextualization within multimodal systems. Against the backdrop of China's cultural "going global" initiative, whether subtitle translation for national image promotional videos achieves cultural expression and international impact has become a key concern in translation studies.

In subtitle translation research, scholars have recognized that subtitle translation is subject to multiple constraints such as temporal limitations, spatial restrictions, and multimodal synergy (Díaz Cintas & Remael, 2007). With the advancement of multimodal discourse analysis (MDA), some studies have begun to explore subtitle translation from the perspective of meaning construction through the interplay of language, images, and sound (Kress & van Leeuwen, 1996; Zhang, 2015). Domestic scholars have also analyzed the multimodal features of subtitle translation in animated films and documentaries based on specific corpora (Yuan & Nai, 2022).

However, existing research remains limited. First, it focuses largely on literary films and narrative documentaries, with insufficient attention to national image-promotional videos. Second, while some studies adopt a multimodal perspective, they lack in-depth discussion of subtitle translation's role in cultural value

reconstruction and international communication, and thus fail to systematically account for its strategic function within multimodal systems. Based on these gaps, employing the international promotional video *CHN* on Chinese civilization as corpus, this study applies Zhang's multimodal discourse analysis framework to conduct a systematic analysis of the Chinese-English subtitle translation from four dimensions. The findings will offer implications for the translation and international dissemination of such audiovisual works, helping to enhance the external communication of Chinese culture.

Literature Review

Multimodal discourse analysis (MDA) is an approach to discourse studies developed on the foundation of social semiotics, emphasizing that meaning is not constructed solely through linguistic signs but through the synergy of multiple semiotic modes such as language, images, sound, movement, and space. Kress and van Leeuwen (1996) argue that non-linguistic resources also possess systematic meaning-making functions and should be regarded as parallel meaning-making resources to language. In audiovisual texts, different modalities collaborate, complement, and constrain one another to co-construct a complete discourse meaning.

In domestic research, drawing on social semiotic theories, Zhang (2015) proposed a relatively systematic comprehensive framework for multimodal discourse analysis, situating multimodal discourse within broader sociocultural contexts. This framework analyzes multimodal discourse across four levels: culture, context, content, and expression, emphasizing the functional division and synergy among different modalities at each level. Specifically, the culture level concerns ideology, values, and genre conventions; the context level examines communicative purposes, participant relationships, and the environment of dissemination; the content level focuses on the construction of conceptual, affective, and interpersonal meanings; and the expression level analyzes the concrete realization of modes such as language, image, and sound.

A Study on Subtitle Translation From the Perspective of MDA

Compared to analyzing subtitle translation solely from a linguistic perspective, MDA offers a more comprehensive lens to reveal the functional positioning of subtitles within audiovisual texts. In this section, the author will analyze the subtitle translation from Zhang's four levels.

Culture Level

The cultural level constitutes the cornerstone of multimodal discourse analysis. Rather than emphasizing superficial cultural comparisons, it focuses on the profound consideration of ideology and genre during the translation process. The translator's fundamental task is to bridge the cognitive gap between source and target cultures, grounded in a thorough understanding of both cultural contexts.

Example 1: 我相信天人合一，万物并育。

Translation 1: I believe that humanity and nature can live in harmony and all living creatures can grow together.

In this example, “天人合一” is a core concept in traditional Chinese philosophy, emphasizing the intrinsic unity between humanity and nature. Instead of employing literal translation, the translator rendered this concept as “live in harmony”, an ecological ethic widely recognized in Western culture. Concurrently, “万物并育” was dynamically interpreted. This translation strategy successfully universalizes distinctive Chinese philosophical thought and circumvents potential comprehension barriers. Consequently, it enhances accessibility among

international audiences with diverse philosophical backgrounds, thereby aligning with the generic requirement of national image promotional films to achieve broad resonance.

Example 2: 我崇尚自强不息，厚德载物。

Translation 2: I advocate for the ceaseless pursuit of self-improvement and embracing the world with virtue.

In this example, “厚德载物” emphasizes accommodating all things through profound virtue. Its translation executes a pivotal transformation: It converts “载物”, a receptive action, into “embracing the world”, a proactive and outward-oriented gesture. This transformation reframes the presentation of moral philosophy, externalizing it from an internal cultivation into a globally-oriented and benevolent action philosophy. It also fulfills the generic objective of promotional films to construct an approachable and proactive civilizational subject.

Contextual Level

The contextual level attends to the specific environment in which discourse occurs. Halliday (1978) proposed the concept of register, comprising three components: field (theme and purpose of communication), tenor (relationship and attitude among participants), and mode (medium and channel of communication). In multimodal translation, translators must accurately apprehend the overall communicative context of the film and ensure that the target text corresponds to this context.

Example 3: 是古老神秘的兵马俑/还是憨态可掬的大熊猫? /是丰富多彩的饮食文化? 还是刚柔并济的中国功夫? 这些答案我曾听过无数次。但我, 远不止于此。

Translation 3: Am I just about the ancient and mysterious terracotta warriors, or beloved and giant pandas? Or do you think of splendid food culture and Kung fu? People often associate me with these things. But I am far more than just that.

This segment establishes the film’s field of discourse: transcending stereotypes to define Chinese civilization multidimensionally. The translation preserves symbolic references through precise correspondence (e.g., terracotta warriors, kung fu) and flexible adaptation (e.g., transforming “I have heard” into “People often associate me with”). Moreover, by incorporating qualifiers such as “just about” and “just that” and using the structure “far more than”, the translation reinforces the core purpose of transcending conventional perceptions, ensuring alignment with the film’s objective in terms of field.

Example 4: 坚守，创新，我一直前进着。

Translation 4: While passing on traditions, I have never ceased to make something new and forge ahead.

This sentence exemplifies adaptation to the mode of discourse. The Chinese source text is highly condensed while the translation enriches the structure using “while” and the present perfect continuous. Particularly noteworthy is the word “forge ahead”, which achieves cross-modal metaphorical synergy. This phrase inherently conveys dynamic imagery of surging upward, establishing direct correspondence with the visual image of fireworks ascending skyward. The physical trajectory of fireworks becomes a vivid visual metonymy for the abstract spirit of “forging ahead”, while the phrase imbues the imagery with connotations of unwavering determination. This resonance between linguistic and visual symbols demonstrates how translation anchors visual meaning through precise lexical choices.

Content Level

The content level focuses on the construction and transmission of meaning itself. According to Halliday (1978), meaning can be categorized into three metafunctions: ideational, interpersonal, and textual. Ideational

meaning refers to the representation of experience and perception of the world; interpersonal meaning concerns the establishment of social relationships; and textual meaning pertains to the organization of discourse to ensure coherence. In multimodal translation, translators must ensure that the target text aligns with meanings conveyed by other modes.

Example 5: 黄河九曲，长江奔流，是我。

Translation 5: The Yellow River winds its way and the Yangtze River rushes across more vast territory.

In terms of ideational function, the Chinese phrase “是我” constitutes a relational process denoting attribution. The translation, however, completely transforms it into material processes, employing two vivid action verbs, “winds its way” and “rushes across” to depict the rivers. This transformation is far from arbitrary; rather, it achieves high coordination with the visual mode: The footage presents dynamic scenes of the Yellow River meandering and the Yangtze River surging forward. Through transitivity system conversion, the translation enables to directly “mimic” the visual dynamics, realizing modal isomorphism in ideational meaning. This allows viewers to perceive more intensely through text the geographical magnificence manifested in the imagery.

Example 6: 多元一体，博采众长，是我。

Translation 6: Embracing unity amid diversity, I’m keen on learning from the strengths of others.

In terms of textual function, the Chinese text positions the theme at the sentence-initial position as the information point of departure, concluding with “是我” as a summary. The translation, however, restructures the thematic organization by placing “I” in the thematic position as the sentence’s point of departure and focal center, while converting the Chinese theme into a present participial phrase “Embracing unity amid diversity” functioning as an accompanying circumstance. The main clause “I’m keen on...” foregrounds the proactive attitude of “I”. This thematic restructuring aligns the information flow with the conventional “subject-action” pattern in English, establishing clear focus (consistently on “I”) and smooth logical progression. It powerfully constructs a civilizational image centered on “I” that actively embodies the principles of inclusiveness and learning.

Expressive Level

The expression constitutes the material dimension through which multimodal meaning is realized. It integrates linguistic signs (e.g., vocabulary, grammar) with non-linguistic signs (e.g., image, sound, layout, typography). In multimodal translation, the translator must possess the ability to weave textual meaning into the entire audiovisual text, ensuring that the translated language synergizes with elements of visual presentation and auditory rhythm.

Example 7: 大漠孤烟，江南细雨，是我。

Translation 7: Wisps of smoke rise in my vast northern deserts while drizzles nurture the prosperous water towns of my south.

This sentence exemplifies profound text-image complementarity. The footage alternately presents the desolate expanse of vast deserts and the misty scenery of graceful water towns. The translation not only faithfully renders the imagery but also elevates mere scenic description into an exposition of “China’s north-south geographical diversity” by incorporating two explanatory geographical-cultural modifiers: “vast northern deserts” and “prosperous water towns of my south”. The linguistic text explicates the underlying cultural-geographical connotations of the visuals, guiding viewers from “observing landscapes” toward “comprehending the spatial configuration of civilization”, thereby achieving semantic expansion and deepening of the visual mode through the linguistic mode.

Conclusion

The present study adopts the perspective of MDA and selects the international promotional video *CHN* on Chinese civilization as a representative case. It conducts an in-depth examination of the characteristic subtitle translations in this video, interpreting the interactive meanings among the visual, auditory, and textual (subtitle) modalities. The findings reveal that its translation practice constitutes a discursive process of meticulously coordinating with the audiovisual modalities across four dimensions—cultural, contextual, content, and expressive—to jointly construct meaning. The research provides critical theoretical insights and methodological implications for understanding and practicing the external narration of China's stories in the new era.

References

- Castro, C. D., & Halliday, M. A. K. (1995). An introduction to functional grammar. *Language*, 71(4), 831.
- Díaz Cintas, J., & Remael, A. (2007). *Audiovisual translation, subtitling*. Manchester: St Jerome Publishing.
- Dong, M., & Yuan, X. L. (2021). A study on the construction of a framework for multimodal aesthetic critical discourse analysis. *Foreign Language Education*, 42(1), 77-82.
- Halliday, M. A. K. (1978). *Language as social semiotic: The social interpretation of language and meaning*. London: Edward Arnold.
- Hu, Z. L. (2007). Multimodality in social semiotic research. *Language Teaching and Linguistic Studies*, 29(1), 1-10.
- Kress, G. R., & Van Leeuwen, T. (1996). *Reading images: The grammar of visual design*. London: Routledge.
- Li, Y. X. (2001). Strategies for subtitle translation. *Chinese Translators Journal*, 23(4), 38-40.
- O'Halloran, K. L. (2004). *Multimodal discourse analysis: Systemic-functional perspectives*. London and New York: Continuum.
- Yuan, X. L., & Nai, R. H. (2022). A study on the meaning construction of multimodal discourse in the international communication of "cultural China". *Foreign Language Education*, 43(5), 23-29.
- Zhang, D. L. (2015). *Multimodal discourse analysis theory and foreign language teaching*. Beijing: Higher Education Press.
- Zhu, Y. S. (2007). The theoretical foundations and research methods of multimodal discourse analysis. *Foreign Language Research*, 30(5), 82-86.