

A Study on the English Translation of ICH Promotional Videos from the Perspective of Multimodal Discourse Analysis

Juan Wu

School of Foreign Languages, University of Sanya, Sanya, Hainan, China

Abstract: The English translation of intangible cultural heritage (ICH) promotional videos has become especially significant in the context of cultural inheritance and international dissemination of ICH. The theory of multimodal discourse analysis provides a multi-layered framework for the translation of ICH promotional videos and serves as a powerful tool for analyzing their translation. This theory highlights the integration of various communication modes such as language, images, and sound, and their crucial role in cross-cultural communication. When translating ICH promotional videos, translators need to keenly capture and balance the different modalities—language, images, and sound—by adjusting the relationships between them to preserve the essence of the original content as much as possible. This paper, based on the theory of multimodal discourse analysis, conducts an in-depth analysis of how multimodal synergy is leveraged in the translation process to enhance the overall effect of ICH promotional video translations. The aim is to offer insights for translating ICH promotional videos, thereby improving the effectiveness of ICH cultural transmission in cross-cultural communication and promoting mutual learning and integration among different cultures.

Keywords: Intangible Cultural Heritage; ICH; Promotional Videos; Translation; Multimodal Discourse Analysis

1. Introduction

In the context of today's globalization, intangible cultural heritage (ICH) carries the unique historical memories, cultural traditions, and spiritual connotations of a nation or region. Its inheritance and international dissemination are of paramount importance. As a precious part of human civilization, ICH is a vivid

manifestation of cultural diversity. It includes a wide range of elements such as folk crafts, traditional skills, folk activities, and oral traditions, encapsulating the wisdom and creativity of our ancestors and bearing witness to the development and changes in various regions over time. However, with the rapid development of modern society and the impact of globalization, many ICH items face the severe challenge of being lost, making the strengthening of ICH preservation and inheritance an urgent task.

Among the various ways of preserving ICH, international dissemination plays an indispensable role. With the deepening of economic and cultural globalization and the rapid development of new technologies such as the internet, artificial intelligence, and the Internet of Things, the demand for global dissemination of ICH is growing, and the means and channels of communication are becoming more diverse. Among these, promotional videos are an important modern multimedia communication tool. As an intuitive and vivid medium, they can enhance the visibility of ICH, foster emotional resonance, and expand the scope of dissemination. They play a crucial role in both the preservation and transmission of ICH. In the cross-cultural communication of ICH worldwide, the translation of promotional videos serves as a vital bridge in intercultural exchanges. The quality of ICH promotional video translations directly affects their dissemination effectiveness. High-quality translations not only spread ICH cultures worldwide, allowing more people to understand and appreciate these precious cultural heritages, but they also help strengthen national pride and cultural identity, inspiring a sense of responsibility and mission for cultural inheritance. Furthermore, such translations contribute to the development of cultural diversity globally, promoting mutual learning and integration between different cultures and

reducing misunderstandings and conflicts arising from cultural differences.

2. The Theory of Multimodal Discourse Analysis and Its Applicability to the Translation of ICH Promotional Videos

Zhang defines multimodal discourse as a communicative phenomenon that uses multiple sensory modes, such as auditory, visual, and tactile, through various means and symbolic resources, including language, images, sounds, and actions [1]. As seen from this definition, multimodal discourse analysis involves the comprehensive use of various methods and symbolic resources, exploring how language and other symbolic resources, such as images, sounds, videos, and bodily movements, interact in communication to convey meaning. Multimodal discourse analysis has not only promoted the development of linguistics and deepened our understanding of semiotics, but has also had a profound impact on fields like aesthetics, literature, rhetoric, and communication studies [2]. Today, it has become an important interdisciplinary research area. Chinese scholars have widely applied in the fields such as advertising [3], film and television [4-5], education [6], and tourism [7]. However, on the other hand, the level of research outcomes in multimodal discourse analysis still requires further improvement [8]. Linguists Gunther Kress and Theo van Leeuwen are widely regarded as pioneers of multimodal discourse analysis. In 1996, they introduced the concept of visual grammar in their book *Reading Images: The Grammar of Visual Design*. Based on Halliday's systemic functional linguistic theory, they extended it to the visual mode, establishing a theoretical framework for visual grammar. They proposed that images have three main functions in communication: representational meaning, interactive meaning, and compositional meaning. This theory has provided a solid theoretical foundation for multimodal analysis [9]. In 2003, Li introduced this theory to China through her paper *Social Semiotic Approach to Multimodal Discourse* [10], marking the beginning of Chinese scholars' research into multimodal discourse analysis.

Later, many Chinese scholars made significant contributions to the development of multimodal discourse analysis in China. Zhang, in particular, focused on localization and,

based on systemic functional linguistics, proposed a comprehensive multimodal discourse analysis framework adapted to Chinese culture and social contexts. He emphasized that multimodal discourse analysis could be studied from cultural, contextual, content, and expressive levels [1].

Multimodal discourse analysis breaks the limitations of traditional discourse analysis, which mainly focuses on the linguistic level. It highlights the interaction between language and other symbolic resources such as images, sounds, videos, etc. in the process of constructing meaning. This comprehensive analytical perspective not only reveals the richness of linguistic expression but also demonstrates cross-media and cross-modal communication methods. ICH promotional videos are a typical example of multimodal discourse, as they often integrate various modalities such as text, sound, images, and video. Through the synergistic effect of these elements, they convey the value and charm of ICH culture to the audience more vividly and intuitively, thus achieving their promotional and educational goals.

The applicability of multimodal discourse analysis to the translation of ICH promotional videos is reflected in several aspects. First, ICH has rich connotations and unique cultural symbols, and it is often difficult to fully and accurately convey its deeper cultural meanings through language alone. Multimodal discourse analysis emphasizes the synergy of different modes, and through the support of other symbolic resources like images and sounds, it can present the uniqueness and depth of ICH more comprehensively, helping the target audience better understand and appreciate these cultural features. By integrating various symbolic resources, translation can not only convey linguistic information but also reflect the emotional and symbolic significance of ICH culture.

Second, there are differences in cultural backgrounds, values, and cognitive patterns between different countries and ethnicities, and when translating ICH promotional videos, it is important to consider how to overcome these cultural differences. Multimodal discourse analysis theory provides translators with a multi-dimensional perspective, helping them better understand the cultural information carried by different modes in the original

promotional video. Based on this understanding, translators can adopt appropriate translation strategies and cultural adjustments to make the translation more in line with the target culture's language habits and cognitive patterns, improving its readability and acceptability and avoiding cultural misunderstandings or misrepresentations. Finally, ICH promotional videos themselves are typical multimodal texts, where elements like images, sounds, and language work together and depend on each other to construct the overall cultural image. Multimodal discourse analysis provides a comprehensive perspective for translating such texts, as translators need to consider not only the linguistic aspects but also the integration and interaction of multimodal elements. For example, when translating an ICH promotional video, how to handle the visual elements (such as the display of traditional crafts), sound effects (such as the emotional rendering of background music), and the complementary role of language is crucial to ensuring translation quality. By considering the coordination of multimodal elements, translators can optimize translation strategies, allowing the translated text to effectively integrate and interact with multimodal elements in the new linguistic and cultural environment, thereby better conveying the essence of ICH culture. In conclusion, the theory of multimodal discourse analysis provides a new approach and methodology for translating ICH promotional videos. By comprehensively considering the synergy between language and other symbolic resources, it can more fully and accurately convey the rich connotations of ICH culture, bridging cultural and linguistic barriers, and promoting the global dissemination of ICH.

3. Introduction to the Promotional Video of the ICH Online Promotion Campaign

The large-scale online promotion campaign for ICH was launched in September 2017. Its aim is to showcase China's significant achievements in the protection of intangible cultural heritage through media platforms such as the internet, television, and newspapers, and to promote the inheritance and development of ICH. Through extensive media coverage, the campaign has attracted widespread public

attention and sparked interest and enthusiasm for Chinese traditional culture.

The promotional video is one of the key components of this campaign, with a duration of approximately 3 minutes and 20 seconds. It highlights the spirit of craftsmanship and showcases the unique charm of traditional skills and culture. The video features both Chinese and English subtitles, enabling a broader global audience to understand and appreciate the value of these precious cultural elements, thereby promoting cross-cultural exchange and integration.

The release of the promotional video has garnered a great deal of positive feedback, including numerous likes and comments from online viewers, and has increased public awareness and support for the protection of ICH. Many viewers expressed pride and admiration for the traditional culture and beautiful landscapes depicted in the video, calling for the continued inheritance and promotion of these cultural treasures. The success of this campaign signifies that the protection of ICH is not only the responsibility of a single country but also a global cultural mission that contributes to the development of cultural diversity worldwide.

4. Multimodal Discourse Analysis on the Subtitle Translation of the ICH Campaign Video

4.1 The Cultural Level

The promotional video for a large-scale online campaign on ICH opens with an inspiring shout, followed by the powerful and eruptive sound of drums. The entire opening sequence is filled with energy and passion. This dynamic rhythm and momentum instantly ignite the audience's enthusiasm, bringing a strong sense of vitality and impact. Against the backdrop of distinctive and powerful drumbeats, eight groups of four-character keywords appear prominently in the center of the screen: "tiān rén hé yī, dào fǎ zì rán, zì qiáng bù xī, hòu dé zài wù, xiū shēn qí jiā, xiào tì zhōng xīn, jìng tiān xī wù, hé ér bù tóng". These opening keywords not only set the cultural tone for the promotional video but also showcase the profound depth of Chinese culture through an expressive linguistic form.

The four-character phrase is a unique linguistic form in Chinese, known for its brevity and

depth of expression. It represents the distinct charm of the Chinese language and its cultural characteristics. In this video, the use of four-character phrases is carefully designed. These phrases are presented in the traditional vertical writing style, appearing one by one in eight frames on the screen, lasting about twelve seconds in total. The rhythm is tight, and the visual impact is striking. However, it is worth noting that these four-character phrases are not provided with corresponding English translations, which raises questions about subtitle translation strategies.

Subtitles are inherently constrained by the need for immediacy and space, especially in fast-paced promotional videos where an overload of text can create a sense of visual clutter and negatively impact the audience's viewing experience. In the brief twelve seconds of this sequence, overly detailed subtitles could make it difficult for viewers to focus on both the text and the imagery, thus diminishing the effectiveness of information transmission. For highly condensed language forms like these four-character phrases, translating them into other languages often requires longer phrases or sentences for clarification, which tends to sacrifice the conciseness and precision of the original expression. To avoid overwhelming the viewer with excessive text, the translators opted for an omission strategy, reducing the cognitive load on the audience and allowing them to focus better on the visuals and emotional impact of the video.

Despite the lack of English translations for the eight four-character phrases, the video does not lose its cultural communicative effectiveness. In fact, it cleverly compensates for the absence of textual translation through the imagery presented in each frame, effectively conveying the cultural connotations of these keywords. For example, the first four-character phrase "tiān rén hé yī", rooted in ancient Chinese philosophical thought, emphasizes the harmonious coexistence of humans and nature, which is one of the core elements of Chinese culture over the past five thousand years. In the *Zhuangzi · On the Equality of Things*, it is said: "Heaven and Earth are born together with me, and all things and I are one." This idea deeply explores the interdependence and symbiosis between humanity and the natural world.

In the promotional video, the visual representation accompanying the first four-character phrase "tiān rén hé yī" presents a vast and tranquil natural landscape: sunlight pours through the clouds and bathes the rolling green mountains, with mist flowing like a veil through the valleys. The mountains and clouds interweave, as if the earth and sky are breathing together. Layered terraces wind down the slopes, illustrating the beautiful harmony between humans and nature. This scene not only highlights the essence of "tiān rén hé yī" but also allows the audience to viscerally experience the harmonious relationship between humanity and nature, a sentiment that is more emotionally resonant than a simple translation of words.

Through this multimodal presentation, the promotional video conveys the cultural essence in a more impactful way, helping the audience understand the deep meaning of Chinese culture and more easily evoke emotional resonance. Therefore, even without the English translations, the promotional video successfully achieves cross-cultural communication and the transmission of values through the combination of images and sound.

4.2 The Contextual Level

In multimodal discourse, contextual factors are determined by the scope, tone, and mode of discourse [1]. In the promotional video, the scope of the discourse focuses on China's intangible cultural heritage, exploring its profound impact and significant value. Through this scope, the video highlights the important role of ICH as a carrier of cultural transmission in modern society. In terms of the tone, the video adopts an impassioned and solemn tone. With highly affirmative language, it conveys a deep sense of pride and identity in the long-standing culture and glorious civilization, emphasizing the crucial role of ICH in contemporary society and expressing a hopeful vision for its future transmission and development. Regarding the mode of discourse, the video uses emotionally charged language, aiming to touch the hearts of the audience. Whether in the expression of language or in the emotional rendering, the video permeates with a deep concern for culture and intangible cultural heritage, demonstrating strong emotional appeal and persuasive power. This emotional penetration not only resonates with

the audience but also encourages viewers to actively participate in the protection and transmission of ICH.

Example 1:

Original text:

Ràng wǒmen yīqǐ qù tàn xún sǎn luò zài zhōnghuá dàdì gègè jiǎoluò de wénhuà zhēnbǎo, qīnshēn gǎnshòu dàibiǎo zhōngguó de “zuìměi fēiyī”, “zuìniú juéhuó”, “zuìchéng jiàngxīn”.

Translation:

Let us explore and seek out the cultural treasures that are scattered in every corner of the Chinese land. Let us see and appreciate firsthand China’s most beautiful intangible heritages, most stupendous skills, and most distinct craftsmanship.

In this example, the original text is full of emotion. The Chinese phrase “Ràng wǒmen yīqǐ...” uses a typical imperative sentence structure, which aims to call upon or mobilize the target group to participate in a particular activity. The phrase carries a clear sense of call and mobilization while also exhibiting an approachable tone. This expression effectively stimulates the desire for collective action. In the translation, the translator has made a moderate addition.

The phrase “Let us explore and seek out...” is added before the original text “qīnshēn gǎnshòu”, thereby further emphasizing the context of the promotional video’s call and mobilization. By using the repetition of “Let us,” the translation creates an atmosphere of collective participation, allowing the audience to more profoundly realize that they are part of the cultural transmission effort. This linguistic strategy effectively strengthens the audience’s sense of belonging and willingness to act, encouraging them to actively engage in the protection and transmission of intangible cultural heritage, ultimately achieving the social impact desired by the promotional video. In this way, the goal of the discourse—stimulating emotional resonance and encouraging participation in the protection and transmission of intangible cultural heritage—is effectively conveyed, further reinforcing the video’s role in transmitting cultural values and promoting social action.

4.3 The Content Level

The content level of multimodal discourse is divided into the semantic level and the formal

level [1]. These two elements are closely connected, jointly constructing the comprehensive communication and in-depth interpretation of information.

Example 2:

Original text:

Zhōnghuá mínzú chuàngzào le yuányuǎnliúcháng de zhōnghuá wénhuà. Zhōnghuá mínzú yě yīdìng nénggòu chuàngzào chū zhōnghuá wénhuà xīn de huīhuáng.

Translation:

The Chinese nation has a rich and long-standing culture, and can certainly bring about a new Chinese cultural beacon.

Example 2 comes from the final section of the promotional video. In the video, the camera focuses on an ancient and thick book, with a page gently turned by the wind. The sound of a crisp page turning follows, and then the voiceover begins. As a vocal modality, the voiceover is often used to enhance the depth of narrative and convey emotions. In this scene, its use is particularly clever, adding more emotional layers and an artistic atmosphere to the message. Here, the visual modality and the vocal modality work together to enrich and complement the meaning conveyed by the textual modality.

In the original text, both sentences have “Zhōnghuá mínzú” (the Chinese nation) as the subject, expressing the brilliance of culture and the hope for the future. The translator appropriately shortened and adjusted the structure, combining the two independent sentences into one parallel sentence, making the phrasing more concise and smoother, while also better suited for the way subtitles are typically displayed. Visually, the first half of the original sentence is paired with images of relics and buildings such as the Terracotta Army and the Temple of Heaven, showcasing China’s long cultural tradition. As the sentence ends, the scene shifts to a group of people playing large horns towards the sky, accompanied by the slowly opening vermilion doors. At this point, the background music is filled with vitality and power, and the imagery and music complement each other, creating a strong visual and auditory impact. The text modality, visual modality, and auditory modality work together in this moment, forming a powerful complement and coordination. Together, they not only convey

the theme of the inheritance and development of ICH but also deepen the emotional and symbolic significance of this theme.

Through this multimodal narrative approach, the audience can feel that the work of protecting and inheriting China's ICH is ushering in a new chapter. At the end of the promotional video, the horn symbolizing the protection and inheritance of ICH has been sounded, and the doors to the world of intangible cultural heritage are slowly opening to the world. The doors in the visual are like a gateway to culture, inviting the audience into this captivating cultural hall. At this moment, the combination of visuals, sound, and text not only allows the audience to perceive the historical depth of Chinese culture but also inspires a sense of responsibility and mission to contribute to the protection and inheritance of intangible cultural heritage. The integration of these three modalities makes the transmission of information more profound and powerful, ensuring that the deep meaning expressed in the promotional video reaches the hearts of every viewer.

4.4 The Expression Level

The expression level mainly refers to the media aspect, encompassing two major categories: linguistic media and non-linguistic media. Linguistic media can be divided into pure linguistic media forms and media forms accompanied by language. Non-linguistic media refer to various non-verbal means used in the process of meaning expression, such as the communicator's body movements and various tools and environments [1]. Linguistic and non-linguistic forms complement each other, forming the core of information transmission.

Example 3

Original text:

Cóng 2001 nián kūnqǔ bèi liánhéguó jiàokēwén zǔzhī lièrù “rénlèi kǒutóu hé fēi wùzhì yíchǎn dàibiǎozuò” dào 2016 nián “èrshìsì jiéqì” de shēnyí chénggōng...

Translation:

From the 2001 proclamation of Kunqu Opera to the 2016 inscribing of China's “24 Solar Terms”...

In the video, example 3 demonstrates China's significant achievements in the protection of ICH through specific examples. When listing concrete instances, the translator employed

moderate omissions, highlighting the two major milestones: the successful application of Kunqu in 2001 and the successful inclusion of the “24 Solar Terms” in 2016. As the subtitles appear, accompanied by passionate background music, the visuals switch to exciting moments of Kunqu performances and captivating images of the 24 Solar Terms song, with the dual stimulation of visual and auditory senses enhancing the transmission of information. This integration of linguistic and non-linguistic media allows the audience to feel the historical depth and cultural value of intangible heritage in a short time.

Overall, the expression level is a crucial part of multimodal communication. Through the organic combination of linguistic and non-linguistic media, the promotional video not only clearly and intuitively conveys the value of cultural heritage but also makes the transmission of information more vivid, emotional, and layered through the interplay of sound and visuals. This approach not only breaks through language barriers but also stimulates cultural identity and emotional resonance among domestic and international audiences through rich audio-visual materials.

5. Conclusion

The English translation of intangible cultural heritage promotional videos is particularly important in the current context of ICH cultural inheritance and international dissemination. Multimodal discourse analysis theory provides a multi-level framework for translating these promotional videos and serves as a powerful tool for analyzing their translation. In terms of cross-cultural communication, multimodal discourse analysis helps to deeply explore the cultural information embedded in various modes such as language, images, sounds, etc. within ICH promotional videos. This enables translators to more accurately grasp cultural connotations during the translation process and adopt appropriate translation strategies to convey these unique cultural elements to foreign audiences.

Regarding the international dissemination of ICH, the effective use of multimodal discourse analysis can better integrate the various symbolic resources in the promotional video, optimize the English translation, and ensure that the translated content retains its

authenticity while also aligning with the acceptance habits and aesthetic preferences of foreign audiences. This allows foreign viewers to understand China's intangible cultural heritage, appreciate its charm, and fosters a deeper understanding of different cultures, thus promoting exchanges and mutual learning among civilizations.

When translating ICH promotional videos, translators should focus on the different modes, such as language, images, and sounds. They should seek a balance among the various symbolic systems and effectively transform and reconstruct these multimodal elements to enhance the effectiveness of ICH promotion in cross-cultural communication.

References

- [1] Zhang, D. On a Synthetic Theoretical Framework for Multimodal Discourse Analysis. *Foreign Languages in China*, 2009(1): 26-32.
- [2] Pan, Y. & Li, Z. A Comprehensive Review of Multimodal Discourse Analysis in China (2003-2017). *Journal of Fujian Normal University*, 2017, (05): 49-59+168-169.
- [3] Li, D. & Chen, J. A Multimodal Discourse Analysis of the Online GIF Poster: A Case Study on One Call Away. *Journal of Shanghai University of Science and Technology*, 2024, 46(03): 211-217.
- [4] Wu, J. & Liu, A. Intertextuality Strategies in the Translation of the Commentary of the National Image Promotion Video "Perspective". *Journal of Xi'an International Studies University*, 2020(2): 22-25.
- [5] Yuan, X. & Nai, R. Research on the Multimodal Discourse Meaning Construction of "Cultural China" in International Communication. *Foreign Language Teaching*, 2022(5): 23-29.
- [6] Zhang, D. The Application of Multimodal Discourse Theory and Media Technology in Foreign Language Teaching. *Foreign Language Teaching*, 2009, 30(04): 15-20.
- [7] Qi, W. Research on the Construction of Eco-Image of Tourism Destinations Based on Multimodal Discourse Analysis. *Journal of Xi'an International Studies University*, 2024, 32(03): 27-32.
- [8] Cheng, R. & Zhang, D. The Current Status, Characteristics, and Development Trends of Multimodal Discourse Analysis in China: Taking Journal Articles as an Example. *Foreign Languages in China*, 2017, 14(03): 36-44.
- [9] Kress, G. R. & Leeuwen, T. van. *Reading Images: The Grammar of Visual Design*. London: Routledge, 1996.
- [10] Li, Z. Social Semiotics Analysis of Multimodal Discourse. *Foreign Language Research*, 2003, (05): 1-8+80.