

# A Study on the Communication Effect of Multimodal Translation of Urban Promotional Videos in the Era of Integrated Media

Min Xu

Chongqing Normal University, Chongqing Municipality, 401331, China

**Abstract:** Against the background of accelerated globalization and technological revolution in the era of integrated media, urban promotional videos, with their visualized and lively characteristics, have become a powerful medium for cities to step onto the global stage. Multi-modal translation, which integrates various symbolic forms, plays a crucial role in breaking through language and cultural barriers and accurately conveying urban images. Therefore, this study analyzes the current situation of multi-modal translation and investigates the communication effect through questionnaires and individual interviews. Based on these findings, it proposes strategies for enhancing multi-modal translation and communication effect of urban promotional videos, aiming to lay a solid theoretical foundation for the effective international dissemination of urban images.

**Keywords:** era of integrated media, urban promotional videos, multi-modal translation, communication effect

## 1. Introduction

As the "visual business card" of a city's image, urban promotional videos are characterized by their brevity, rich content and convenience of dissemination, and can effectively and intuitively reflect the core ideas of the "city brand" [1]. With the comprehensive infiltration of new media and digital technology, the information dissemination model has been completely revolutionized in the era of integrated media. As a powerful weapon for shaping the city brand, the dissemination scope and influence of urban promotional videos have grown exponentially. However, in the process of cross-cultural communication, language and cultural differences have become huge obstacles between the city image and international audiences. As an innovative translation paradigm, multi-modal translation brings hope to break this predicament. Currently, academic exploration of its communication effect in the field of urban promotional videos is still in the exploratory stage. In-depth analysis of multi-modal translation in urban promotional videos has far-reaching practical significance for enhancing the international image of cities and strengthening their soft power.

## 2. Analysis of Current Situation of Multi-modal Translation and Dissemination of Urban Image Promotional Videos

Integrated media utilizes the Internet as a carrier to comprehensively integrate various media such as radios, televisions, newspapers, and mobile media in terms of human resources, news gathering and editing, interaction, and marketing, thus achieving the goal of "integration of resources, compatibility in content and mutual benefit" [2]. The transformation of traditional media to integrated media has provided a broader space and platform for the production and dissemination of urban promotional videos, but in the meantime posed greater challenges to the translation and dissemination of subtitles in these videos.

## **2.1 Development of Urban Promotional Videos in the Era of Integrated Media**

In the era of integrated media, the development of urban promotional videos are rapid and diverse. In terms of content, they no longer focus merely on surface-level landscapes but delve into the city's essence, integrating historical heritage, cultural customs, and innovative vitality. In terms of form, they shift from traditional linear narratives to diverse interactive formats, using technologies such as VR and AR to create immersive experiences. The dissemination platforms are also increasingly diverse, with social media and short-video platforms becoming important channels, enabling the videos to precisely reach different audience groups and greatly expanding the breadth and depth of dissemination, injecting new vitality into the promotion of urban images.

## **2.2 Application Forms of Multi-modal Translation in Urban Promotional Videos**

In urban promotional videos, multi-modal translation plays a crucial role in promoting the city's image through various forms. In the linguistic modality, accurate subtitle translation bridges the language gap, while voice-over adds emotional warmth to the text. In the visual modality, images of iconic buildings and cultural scenes directly showcase the city's appearance, complementing the textual information. In the auditory modality, background music and the sounds of the city create a unique atmosphere. These modalities interweave, for instance, when presenting traditional festivals. Subtitles introduce customs, pictures depict celebrations, and music sets the mood, thus comprehensively and stereoscopically conveying the city's charm.

## **2.3 Current Problems in Multi-modal Translation of Urban Promotional Videos**

Currently, there are some problems in multi-modal translation for the subtitles of urban promotional videos. In terms of quality, some translated versions of subtitles contain grammatical errors and inappropriate word usage, affecting the accurate transmission of information. In terms of modality coordination, the synchronization between pictures and voice-over is not always accurate, leading to confusion in information delivery. Regarding audience adaptability, there is a lack of in-depth consideration for audiences from different cultural backgrounds, making it difficult for the translation content and presentation forms to resonate widely and therefore weakening the dissemination effect. Traditional translation studies have focused mainly on the language of the translated text, neglecting other aspects of information transmission. With the advent of the era of integrated media, the translation of subtitles in promotional videos is no longer limited to the transmission of text information but presents a system that encompasses the association of words, images and voice, etc [3].

To address these issues, it is necessary to conduct in-depth research on the dissemination effect of multi-modal translation in urban promotional videos in the era of integrated media, thereby providing guidance for the future translation and dissemination of such videos.

## **3. Research on the Dissemination Effect of Multi-modal Translation in Urban Promotional Videos**

This study intends to combine questionnaire surveys and individual interviews to conduct in-depth investigations and interviews with domestic and international audiences of the promotional videos of the Chengdu-Chongqing Twin-City Economic Circle, extracting effective strategies to enhance the dissemination effect of multi-modal translation.

### **3.1 Research Methods**

This study employs questionnaire surveys and individual interviews. The questionnaire survey method involves collecting relevant data through written forms with strictly designed questions to investigate the research subjects. [4] The researchers distributed questionnaires to the target audience who have lived abroad for a long time and have a preliminary or considerable understanding of the Chengdu-Chongqing Twin-City Economic Circle. After sharing the relevant videos through online links for the target audience to watch, they were asked to fill out the questionnaire on the communication effect of multi-modal translation of urban image promotional videos in the era of integrated media. After receiving a certain number of responses, the results were statistically analyzed and objectively and scientifically evaluated.

Based on this questionnaire survey, to make the survey results more reliable and convincing, the researchers also conducted individual interviews with the respondents through online chat to obtain more in-depth insights and understanding.

### **3.2 Research Process**

The process of the questionnaire survey mainly includes the design, distribution, collection, statistics and analysis of

the questionnaire. Multi-modal translation involves multiple aspects such as text, image and sound, so the questionnaire design needs to cover the influence of these elements on the audience's cognition, emotion and behavior. In the design of questions, first, it is necessary to ensure that the questions cover all modalities, such as visual, auditory and textual modalities, as well as their synergy. Second, cultural differences of the foreign audience should be considered, and suggestive questions should be avoided to ensure the neutrality of the questions. The third point to pay attention to is the quantification of communication effect, such as the increase in interest, willingness to share and travel after watching the video. The questionnaire will use the Likert scale to measure these aspects. Under the guidance of the above design ideas, the questionnaire is ultimately divided into the following five parts: basic information of the respondents, evaluation of multi-modal translation effect, cultural adaptability, communication effect, and open feedback. After comparing and considering the urban image videos of the Chengdu-Chongqing Twin-City Economic Circle, among many foreign promotional videos of cities, the author chose the YouTube version of the Chengdu image promotional video: Chengdu - The City You Never Want to Leave. This promotional video depicts and promotes the life of Chengdu from five aspects: eating, drinking, playing, and wearing. The subtitles are in Chinese-English bilingual format, with concise and clear translation. The pictures skillfully switch between traditional Chinese elements such as Hanfu and modern urban elements, complementing each other. The English commentary has a natural and powerful pronunciation, and is accompanied by a lively, energetic and culturally distinctive voice-over. The entire promotional video achieves a harmonious unity in image, text and sound, which is closely related to the multi-modal translation adopted. A total of 100 questionnaires were distributed through email and foreign websites (such as YouTube), and 80 valid responses were received. The results of these 80 questionnaires were analyzed with the assistance of AI. To ensure the reliability of the questionnaire survey results, the researchers also conducted close individual interviews with five respondents to ensure the validity of the research results. The following is a summary and analysis of the key issues involved in the survey and individual interviews.

The core part of the questionnaire includes technical dimension, cognitive dimension, emotional dimension, prediction of communication behavior and direction for optimization of multi-modal translation. From the technical dimension, most of the respondents believe that the subtitle translation of the promotional video meets the basic quality standards, but they have doubts about the handling of cultural-specific items. Among them, 72% of the respondents believe that the subtitle translation of the promotional video is basically accurate, and 38% of the respondents think that "Du Fu" should be translated with annotation. Regarding the synergy advantage of modalities, 61% of the respondents chose [subtitles + close-up of food] as the best combination, demonstrating the complementary and enhancing effect of textual and visual symbols. One aspect worthy of particular attention in the cognitive dimension is the investigation of the multi-modal cognitive gradient, which shows that those with no knowledge of Chinese rely more on [narration + background music] (67%), while native speakers pay more attention to [interviews + scene transitions]. The most directly related factor to the effect of multi-modal translation and dissemination is the prediction of dissemination behavior. The survey results show that the potential for fragmented dissemination on social media is huge. Among them, 58% of the audience choose to "like after watching the clips", which is mutually verified by the average viewing time of 7.2 minutes (platform data). In terms of interest conversion directions, tourism (76%) and cultural studies (43%) are the top two interest areas, but the conversion rate for business investment is only 12%. Regarding the optimization direction of multi-modal translation, 43% of the audience choose cultural annotations for subtitles, and 38% choose to match the music rhythm.

In the individual interview section, the researchers mainly selected five representatives based on the Chinese language proficiency of the interviewees, including an intermediate-level participant with an HSK4 certificate, a native speaker at an advanced level, a beginner with no prior knowledge of Chinese, a beginner with an HSK2 certificate, and an advanced participant with a master's degree. Through the interviews, some common viewpoints were discovered: 1. The problem of spatio-temporal dislocation: Three people pointed out the contradiction between the speed of scene transitions and the duration of subtitles (average reading speed of 0.4 seconds per word vs. actual duration of 0.3 seconds); 2. The expectation of cultural decoding: Four people hoped that important cultural symbols would be translated in a "term + emoji annotation" format that is friendly to social media; 3. multi-modal interference: All respondents mentioned at least one case of mutual exclusion among sound, image, and text (such as traditional musical accompaniment with skyscraper scenes).

#### **4. Strategies for Improving Multi-modal Translation and Dissemination of Urban Promotional Videos**

Based on the combination of the above questionnaire survey and individual interviews, to effectively enhance the dissemination effect, it is necessary to explore practical methods from the dimensions of translation strategies, modal coordination, channel expansion, and audience positioning.

##### **4.1 Optimize Translation Strategies**

Optimizing translation strategies requires translators to have a deep understanding of the content of urban promotional videos. When dealing with culturally rich expressions, they should not simply translate them directly. For example, when "Du Fu" is mentioned in the video, it should be explained as "Du Fu, a famous poet in Tang Dynasty" so as to make the target audience understand more clearly. Additionally, the translation team should keep up with language development trends and use vocabulary and expressions familiar to the target audience to accurately convey the city's unique features, ensuring that the translation text retains cultural charm while conforming to the language habits of the audience, thereby enhancing the dissemination effect.

##### **4.2 Strengthen Modal Coordination**

Strengthening modal coordination requires efforts throughout the production process of the promotional video. During the preparation stage, directors, copywriters, visual designers, and other personnel should jointly discuss and determine the core elements of each modality based on the city's image positioning. During the production process, the time axes of each modality should be strictly calibrated to ensure that scene transitions are synchronized with narration and subtitles, presenting the city's image in a comprehensive and vivid manner and enhancing the video's appeal. During the translation process, translators need to comprehensively consider and strive to achieve the integration of text and visual modality information, exerting a relatively significant impact on the audience through the manipulation of text modality symbols, thereby achieving the communication purpose of the promotional video in conveying information and highlighting the image<sup>[5]</sup>.

##### **4.3 Expand Dissemination Channels**

To expand the dissemination channels, it is necessary to fully utilize the diversity of the new media ecosystem. For short-video platforms, it is advisable to produce 15-60 second clips that are compact in plot and topic-driven, and leverage the popular tags and algorithm recommendations of the platform to attract a large number of users to pay attention and share, such as creating a series of short videos with the theme of "Hidden Gems of the City". On professional video platforms, high-definition, well-produced and detailed complete promotional videos can be released to meet the needs of professionals and deep enthusiasts for in-depth exploration of the city's image. At the same time, cooperation with local TV stations can be sought to broadcast promotional videos during prime time, leveraging the credibility and wide audience base of traditional media to expand the coverage of dissemination and increase the exposure of the city's image through multiple channels.

##### **4.4 Precise Audience Targeting**

Precise audience targeting requires comprehensive research and analysis. For Western audiences, the promotional videos of cities should highlight the modern and international aspects of the city, such as introducing the city's international business center and cutting-edge technological achievements. For Eastern audiences, we should delve into traditional cultural heritage, like telling the ancient legends and traditional customs of the city. Also, language proficiency should be considered. For non-native English speakers with lower language proficiency, the translator needs to provide simple and straightforward subtitle translations and pair them with vivid images to facilitate understanding. At the same time, based on age differences, trendy and fashionable elements should be incorporated for the younger group, such as city music festivals and creative districts, while for the elder group, historical inheritance and humanistic care ought to be highlighted. Through precise customization, the recognition of a city's image among different audiences can be enhanced.

#### **5. Conclusion**

The research on the dissemination effect of multi-modal translation of urban promotional videos is of great significance in the era of integrated media. With the solid proof obtained from questionnaire survey and individual interview, this paper provides a more objective description of the communication effect of multi-modal translation of promotional videos for city image. By optimizing translation strategies, strengthening modal coordination, expanding

dissemination channels, and precisely targeting audiences, the existing problems in multi-modal translation can be effectively overcome. This not only helps to improve the dissemination quality of urban promotional videos but also comprehensively shapes the unique charm of the city and promotes its broader and deeper image dissemination on the international stage.

### **References**

- [1] Huang Shujun. The Translation of City Promotional Video Based on Multi-modal Discourse Analysis ——A Case Study of Hangzhou Promotional Video for G20 Summit [J]. Journal of Ningbo Institute of Education, 2017(10): 94-98.
- [2] Xu Min, Long Minghui. Multi-modal Translation of City Promotional Videos from the Perspective of Communication Studies[J]. Modern English, 2023(22): 80-85.
- [3] Liu Bin, Jia Xue. The Commentary Translation Study of Province Presenting Film from the Perspective of Multi-modal Discourse Analysis [J]. Journal of University of South China, 2022(2): 102-108.
- [4] Chen Jianlin. Modern Foreign Language Teaching and Research: Theory and Method [M]. 2004.
- [5] Ji Wenkai. Multi-modal Coordinated Restructuring in Translation of Culture-introducing Videos [J]. Foreign Language Education, 2021(9): 82-86.

### **Acknowledgments**

This paper is supported by one of Chongqing's Social Sciences Program for Societal Associations (2023SZ23) entitled "Research on Multi-modal Translation & Dissemination of Urban Promotional Videos in Chongqing-Chengdu Economic Circle in Integrated Media Era".