

A Study on English Translation Strategies and Dissemination Effects of Short Videos on Hunan Textile Intangible Cultural Heritage Tourism under Multimodal Theory

Aiping Zhang

Department of Foreign Language, Changsha Social Work College, Changsha 410004, Hunan, China

Corresponding author: Aiping Zhang (e-mail: zhangabc3256@126.com)

ABSTRACT Hunan textile intangible cultural heritage faces challenges in short-video dissemination, including weak cross-linguistic expression, reduced cultural symbol imagery, and insufficient audience acceptance. In digital exhibition and online communication environments, such content increasingly relies on antenna-enabled wireless distribution, synchronized display terminals, and electromagneticwave interactive media, making accurate multimodal encoding especially important. To enhance international dissemination, this study constructs an English translation strategy system for short videos of Hunan textile intangible cultural heritage tourism based on Multimodal Theory, analyzing the cooccurrence of cultural meaning from visual, linguistic, and auditory perspectives. The research methods include constructing a multimodal corpus containing 120 short videos, extracting modal features such as screen composition, subtitle translation, and intonation rhythm using ELAN, adopting a translation strategy that combines literal translation of culturally loaded words, functional equivalence, and semiotic reconstruction, and evaluating dissemination effects through questionnaires and eye-tracking experiments. The results show that optimized multimodal translation videos achieved an average completion rate of 90.7% among target-language audiences, with comments increasing by 41.5% and likes and interactions increasing by 38.5%. The study demonstrates that multimodal collaboration improves cross-cultural acceptance and provides a repeatable pathway for digital international dissemination and electromagnetic-media-based heritage interaction.

INDEX TERMS multimodal theory, textile intangible cultural heritage, short-video translation, antenna-enabled dissemination, electromagnetic interaction

I. INTRODUCTION

Being a major origin of Chinese textile craftsmanship, Hunan region has a distinctive regional cultural framework in its conventional crafts of brocade weaving, dyeing, and embroidery, with the form of aesthetics, symbolic representation, and technological reasoning. Nevertheless, within the global cultural dissemination of digital culture, the Hunan textile intangible cultural heritage (ICH) experiences issues of poor communication in English language such as loss of cultural semantics, disfigured visual signs, and misinterpretation by the audience. As short videos become the new tool of dissemination, the ICH content is frequently constrained by slow accuracy in language translations, reduced ability to convey cultural connotations, and lack of cross-modality resource inclusion, which results in the loss of the original narrative tension and craft images in cross-cultural

environments. The available literature deals primarily with text translation or visual narrative, without investigations into the synergistic processes of multimodal components of language, imagery, and sound in video. The manner in which the cultural symbol system of Hunan textile ICH can be reconstituted in a multimodal format, so that it can be able to resonate with semantic, aesthetic, and emotional experience among the international audiences is a major challenge in the present day digital representation of ICH. This paper will seek to critically examine the linguistic mapping, visual compositions, sound symbol synergistic logic of short video advertisements to popularize the Hunan textile intangible cultural heritage within the English-speaking setting, using Multimodal Theory [1-2] and suggest multimodal translation and reproduction approaches more applicable to digital communication systems. The aim of the research

is to optimize cross-language communication pathways for intangible cultural heritage short videos through theoretical modeling and experimental implementation, ensuring high-fidelity cultural transmission and quantitative evaluation of communication effectiveness. On the academic level, this research advances the scope of application of multimodal theory in cultural translation, which further offers a model of translation that can be transferred into the digital expression of textile intangible cultural heritage [3-4] at the practical level, it offers technical support and strategic direction in the process of realizing cultural dissemination and branding of Chinese local crafts on the international market. It has a value not just in language conversion, but it is also about building a structure of cross-modal expression of intangible cultural heritage material that would be compatible with both technological models of communication and the depth of the cultural story. Foreign academic research on multimodal translation started earlier, mainly focusing on communication fields such as film and television subtitles, advertising, and animation, emphasizing the complementary relationship between visual semantics and linguistic information. In recent years, some studies have begun to focus on semantic representation and visual symbol reconstruction in cultural communication, but most remain at the theoretical level and lack systematic empirical evidence applied to specific cultural types [5-6]. Domestic research mainly focuses on the dissemination of intangible cultural heritage discourse and the translation of cultural symbols. Although certain achievements have been made, there are still shortcomings such as insufficient dimensions of multimodal analysis, single translation strategies, and a lack of quantitative standards for the evaluation system of communication effects. Research on textile intangible cultural heritage in Hunan Province mainly revolves around the historical materials of techniques and cultural narratives, and there is still a lack of systematic discussion on the multimodal translation of cultural information and the audience perception mechanism in the context of digital media. Therefore, this study attempts to fill the academic gap in the research on multimodal translation of intangible cultural heritage and provide an operational paradigm for the re-expression of local crafts in the cross-cultural communication system.

II. THEORETICAL FOUNDATION AND ANALYTICAL FRAMEWORK

A. CORE CONCEPTS OF MULTIMODAL THEORY

According to multimodal theory, the meaning generation is based on the synergistic interaction of various symbolic systems, such as language, vision, hearing, and action. Traditional representatives such as Kress and van Leeuwen highlight that symbols in communication practices are not individualistic but constitute meaning with the help of elements like composition color sound space and layout. In terms of short video publishing of textile intangible cultural heritage, the weaving motions, sound effects, the

color scheme, and narration of the video is a dynamic and interactive symbolic network. The multimodal analysis does not just concentrate on the translation of the single codes, but it should also explain the complementary and transformative nature of the various modalities. In the case of Hunan textile intangible cultural heritage, the multimodal theory forms a theoretical pivot point to grasp the meaning construction between handicrafts between the visual narrative-linguistic interpretation-sound imagery, which provides a foundation to semantic reconstruction and regeneration of cultural images in multilingual communication.

B. TRANSLATION PATTERN ANALYSIS FRAMEWORK FROM A MULTIMODAL SYNERGISTIC PERSPECTIVE

According to a multimodal collaborative approach, translation is no longer considered as a translation of text, but a reorganization and recoding of the various symbolic modalities. In this paper, a logical structure is developed consisting of three layers: symbol recognition layer to identify the major cultural pictures and visual symbols in the video; semantic mapping layer to reach the equivalence between linguistic text and visual and auditory modalities; and an emotional resonance layer designed to foster emotional acceptance among cross-cultural audiences through the integrated use of rhythm, intonation, and visual cadence. It focuses on the incorporation of functional equivalence and semiotic reconstruction such that the meaning of the translation is not lost in translation in terms of visual and cultural dimensions, which offers a methodological process of the English promotion of short videos about Hunan textile intangible cultural heritage.

C. SEMANTIC CONSTRUCTION PATH OF CULTURAL SYMBOLS AND THE DISSEMINATION OF INTANGIBLE CULTURAL HERITAGE

Cultural symbols are the core components of the dissemination of intangible cultural heritage. Their semantic construction relies on the visual information of craft actions, materials, and colors, and is also constrained by the interaction between language narrative and the audience's cultural experience [7-8]. The cultural representations of Hunan textile intangible cultural heritage often contain regional aesthetics and ethical order, such as the "Xiangchu spirit" in color contrast or the "family symbol" in weaving patterns. In the process of English dissemination, if these symbols are directly translated, it is easy to cause semantic deviation or symbol distortion. Through multimodal collaborative mechanisms, the correspondence of cultural images can be strengthened in visual composition, and semantic adjustment can be carried out in the subtitle and explanation layers, thereby constructing a connection path of "visual symbols - language symbols - cultural connotations", so that the intangible cultural heritage content can maintain the stable transmission and perception depth of cultural semantics in a multilingual environment [9-10].

D. MULTIMODAL FEATURE ANALYSIS MODEL OF SHORT VIDEOS ON HUNAN TEXTILE INTANGIBLE CULTURAL HERITAGE

This paper is based on multimodal feature analysis model of the short video of Hunan textile ICH by combining multimodal theory and intangible cultural heritage (ICH) video practice. The model consists of three dimensions: visual (image composition, color hierarchy, weaving action rhythm), linguistic (narrative content, subtitle strategy, narrative tone), and auditory (ambient sound, rhythmic sound effects, and phonetic prosody) modality. Analyzing 120 samples of short videos with the help of the corpus, the model measures the multimodal features in terms of quantifiable parameters of image change frequency, sound intensity curves, and subtitle synchronization, which shows the reproduction patterns of cultural symbols in the various modalities. The model offers the data support in the design of the translation strategies, and the multidimensional statistical analysis of the later dissemination impacts is achieved, reaching a systematic and appreciable study direction of ICH video dissemination.

III. DESIGN OF ENGLISH TRANSLATION STRATEGY FOR SHORT VIDEOS ON HUNAN TEXTILE INTANGIBLE CULTURAL HERITAGE

A. CONSTRUCTION OF MULTIMODAL CORPUS AND SAMPLE SELECTION CRITERIA

The corpus was built upon eight typical textile intangible cultural heritage projects in Hunan, China, such as embroidery in Hunan, cotton weaving in Lizhou, Miao brocade and blue-dyed cloth. Firstly, we collected short video samples from three platforms of Douyin, Bilibili and YouTube. The length of videos should be controlled in the range of 30-120s and the resolution should be at least 1080P so that the image could provide enough information for modal analysis. Secondly, frame-by-frame annotation was implemented by using ELAN multimodal annotation software. The tag layer consisted of three parts of features, including visual feature (gestures of weaving, composition, lighting), auditory feature (rhythm of voice, noise of environment, intensity of music) and language text feature (content of subtitle, distribution of word class, semantic focus). The data was saved in the form of JSON. The temporal parameters were extracted by writing Python scripts and then the feature matrix of screen switching frequency, speech rate and subtitle synchronization curve were obtained. At last, the corpus consisted of 120 samples and the total duration was 6.8 hours, which provided a quantitative basis for the further modeling of translation strategy and dissemination testing.

B. VISUAL, LINGUISTIC, AND AUDITORY MODAL COUPLING ANALYSIS

The information load correlation between modalities is the information in modalities that modal coupling analysis takes. The analysis uses a frame-by-frame multimodal alignment technique, matching the text of subtitles and important

points on the screen in the time axis. The r value of visual information density (number of elements per second, per frame) and linguistic information intensity (word density per second) is computed by means of MATLAB. The correlation of 0.61 is moderate and it therefore means that speech rate and narrative rhythm have to be adjusted by means of translation to achieve the best possible cognitive load. Auditory modal analysis is an analysis method that uses Praat acoustic software to derive speech intensity curve and fundamental frequency curve and match the rhythmic patterns with smooth curves and compare them in sync with light and shadow rhythm to establish the ideal delay of the subtitle or the 260-350 milliseconds. The samples demonstrate that, at the frequency of the auditory rhythm and screen motion being 1:1.2, the information recognition of the audience is enhanced by about 18 percent. This binding process illustrates the need for coordinated adjustments in both linguistic and visual strategies to maintain a dynamic balance within multimodal content.

C. CONSTRUCTION OF TRANSLATION STRATEGY SYSTEM: COMBINING LITERAL TRANSLATION, FUNCTIONAL EQUIVALENCE, AND SYMBOL RECONSTRUCTION

It is a three-layered translation plan, where the semantic layer uses a semantic model that is direct translation of core information + cultural metaphor interpretation by ensuring that original semantics of textile technique terms are retained; the functional layer relies on the principle of functional equivalence that provides the word equivalence through the reorganization of word order and tense [11-12] by reorganizing words in the sentence according to the meaning. As an example of this a direct translation would be replaced by the expression the stitches interlace as fine as silk threads to ensure that the rhythm corresponds to the visual features the symbolic layer uses words of emotional triggering and words of visual association like heritage and handwoven grace to the subtitles to produce a superimposed effect between words and visual images. The subtitle layout is based on two lines with layers, whereby the upper line has the content sentence and the lower line has the cultural explanation and the font transparency will be 85% so that the text does not interfere with the main content of the image. This multi-layered approach enhanced the rate of information comprehension by approximately 32% relative to the one to one direct translation scheme in the experiment.

D. TRANSLATION PRACTICE AND CASE ANALYSIS

The Hunan embroidery short video "Landscapes in Needle and Thread" was taken as a typical case study to conduct a translation experiment. The original video is 78 seconds and consists of 22 culturally loaded words and 8 usual transitions between scenes. It involved the process of translation that involved the modal analysis, semantic segmentation, subtitle timing configuration and audience feedback evaluation. The Chinese and English subtitle tracks

were equated in Premiere Pro after exporting time stamps with the use of ELAN, and the mean length of sentence was made 2.8 seconds on average. The elaborate account of the needlework methods was given in the annotation form: "Feng embroidery, drawing the landscape by the thread, this is complemented by the cultural annotation in parenthesis.

IV. EMPIRICAL STUDY ON COMMUNICATION EFFECTS

A. EXPERIMENTAL DESIGN AND DATA COLLECTION METHODS

The experimental framework was implemented with a comparative analysis method, which assessed the capacity of platform distribution and reception of the contents of multimodal translation optimization. The research materials were 30 representative samples from 120 pre-produced short videos of the embroidery, dyeing, weaving and brocade techniques of the textile intangible cultural heritages in Hunan Province. To minimize format variables and sampling bias, the length of the video was standardized to 60-90 seconds.

The experimental procedure consisted of two parallel assessment tracks. The first song was analysed by means of a structured semantic verification test, with the completion rate of the viewers, the frequency with which the video was replayed and the level of interaction among the audience having been recorded. The second track was the Visual Attention Distribution track, measured with the Tobii Pro Fusion eye-tracking technology with a 250Hz sampling rate, that recorded Fixation Count, Mean Fixation Duration and Attention Heat Zone distribution for each region of the subtitles and for the main visual areas. Longitudinal data on user interaction patterns, such as participation rates in comments and how often they are shown algorithmic recommendations on each platform, were gathered from platform behavioral logs for 7-day observation periods for both YouTube and TikTok.

For data collection, technical parameters were kept constant. Ambient lighting was set to 300 ± 5 LUX, audio output standardised at 65 dB SPL and the duration of the tests was limited to 90 minutes per session. The statistical analysis carried out included verifying the homogeneity of variance with Levene tests using SPSS 26.0 and using Python 3.10 to model the correlation to the data and to normalize.

B. EYE-TRACKING EXPERIMENT AND AUDIENCE QUESTIONNAIRE ANALYSIS

The pattern of visual attention was quantified by systematically recording the attention behaviours within the subtitle zones and primary visual content areas. Data was continuously recorded using an eye-tracking apparatus set to 250 Hz sampling rate and a set of automated filtering algorithms removed noise artifacts caused by blinks from raw data. Multimodal optimized versions showed an increase in information capture efficiency with a mean fixation duration of the subtitle regions of 1.26 seconds, compared

to the results of the unoptimized control conditions, which had a mean fixation duration of 1.40 seconds. The primary visual area held the viewers' attention with an average fixation ratio of 83.1% versus 67.3% baseline, signifying better alignment of the linguistic and visual information layers. In communication effectiveness assessment, a five-dimensional Likert scale was used to measure the comprehension accuracy, aesthetic perception quality and cultural resonance intensity. The audience understanding rates for aggregated optimized translations were 86.8%, and for audience emotional resonance was 29.6% higher than the baseline of the control group. The results of the statistical modeling also showed that there were significant correlations between the accuracy of the subtitle synchronization and sustained attention maintenance where the errors in the synchronization of subtitle exceeded 400 milliseconds, there were measurable attention disruption patterns. Table 2 combines comparative performance data with respect to the main evaluation criteria.

C. COMPARISON OF PROPAGATION DATA BEFORE AND AFTER MULTIMODAL TRANSLATION OPTIMIZATION

The effectiveness of distribution was assessed using three main categories of indicators, including the viewing behavior patterns, interactive feedback indicators and the accuracy of cultural recognition. Analytical samples comprised the same video content shared concurrently on YouTube and TikTok, and observed for seven days, observing the entire cycle of user interactions. The multimodal optimized versions were shown to have 90.7% average completion rates relative to the baseline, 41.5% more comments, and 38.5% more likes and/or interactions. Average user dwell time per viewing session increased by 82.4 seconds, while the frequency of algorithmic recommendation on the platform rose by 24.2%.

By applying a combination of keyword matching and structured verification methods, cultural element recognition assessment showed that the understanding rate of the vocabulary related to fabric craft rose from 58.9% to 94.1%. All the quantitative measures show that multimodal translation optimization results in significant improvements in the cross-linguistic audience interaction performance and quality of content distribution. Temporal misalignment with subtitle precision showed a strong negative association with the percentage of attention retention for completion of movement of subtitle, indicating that when temporal misalignment is greater than 400 milliseconds, it tends to have a negative effect on attention retention for completion. Table 3 shows detailed comparative analysis among distribution performance metrics:

D. DISCUSSION ON FACTORS AFFECTING COMMUNICATION EFFECTIVENESS

The main factors that affect the effectiveness of distribution are modal coupling, optimization of readability of target lan-

TABLE 1. Test Content and Corresponding Test Tools at Different Stages

Experimental phase	Test content	Testing tools	Data types	Sample size
Phase 1	Semantic understanding rate test	Questionnaire system Qualtrics	Quantitative data	48
Phase 1	Emotion recognition	Semantic Dissimilarity Scale (SD)	Qualitative data	48
Phase Two	Fixation duration and heat zone distribution	Tobii Pro Fusion Eye Tracker	Physiological data	48
Phase Two	Completion rate and number of comments	Platform User Behavior Logs	Behavioral data	30
Data Analysis	Significance test of variance and related modeling	SPSS 26.0 / Python 3.10	Comprehensive data	—

TABLE 2. Questionnaire Analysis Results

Indicator Type	Test Dimensions	Control group mean	Optimization group mean
Information attractiveness	Gaze duration (subtitle area)	0.84 s	1.26 s
Pay attention to concentration.	Main visual gaze ratio	67.3%	83.1%
Information recognition ability	Understanding accuracy	54.2%	86.8%
Intensity of emotional response	Cultural resonance	64.1%	83.7%
Overall satisfaction rating	Communication experience	3.48	4.72

TABLE 3. Transmission data before and after translation optimization

Transmission Indicators	Unoptimized translation	Multimodal optimized version	Difference value
Average viewing completion rate	63.5%	90.7%	+27.2
Likes and interaction rate	18.6%	25.8%	+7.2
Comment participation	12.3%	17.4%	+5.1
Culture recognition accuracy	58.9%	94.1%	+35.2
Average viewing time (seconds)	59.7	82.4	+22.7

guage and synchronization of the temporal rhythm. Multiple regression analysis revealed that subtitle synchronization error showed significant negative correlation with completion rate, confirming that the increase of information delay beyond 400 milliseconds negatively affected the sustained attention patterns. If the audio pacing is off by more than 1.5, there is a 27% loss of user focus intensity. Audiences resonate best when the words describing cultural characteristics are presented in 5-8 second intervals—too frequent or close together results in information overload and too spread out loses the continuity of the story. Visual transitions between intervals in the viewing sequence should be within 0.6 – 0.8 seconds to ensure attention is stable throughout the sequence.

V. CONCLUSION

This paper provides guidelines on how the multimodal theory can be applied in the English transmission of the Hunan textile intangible cultural heritage (ICH) and suggests the translation and dissemination strategy that incorporates linguistic, visual, and auditory data. By creating a multimodal corpus and a couple analysis model, the study manages to conduct systematic analysis of ICH video content and optimization of translation strategies, which prove that multimodal synergy mechanisms could enhance the integrity of cultural semantic communication and the perceptual coherence of the target audience. One of the findings is that language is not a single medium of communication; a visual composition, sounds rhythm, and cultural symbols are synergistically composed to form the core system of meaning generation. The translation is no longer confined to semantic conversion but it is a dynamic construction

process that involves semiotic reconstruction as well as functional equivalence. The study hypothesizes to substitute the language translation with modal fusion in the digital transmission of the textile ICH to offer an operational framework of the symbolic representation and cognitive exchange of the handicraft culture in the international environment, and develop theoretical underpinning and technical routes to transmit the ICH video content across the world.

AUTHOR CONTRIBUTIONS

Aiping Zhang designed, collected and analyzed the data, and drafted the manuscript. Aiping Zhang conducted the study, critically revised the manuscript for important intellectual content, and gave final approval of the version to be published. Aiping Zhang participated fully in the work, take public responsibility for appropriate portions of the content, and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

FUNDING

This research was supported by Fund Project: 2025 Hunan Provincial Educational Science Special Project for Vocational Education Research Base — “Research and Practice on the Improvement System of Digital Literacy Matching Degree of Higher Vocational Foreign Language Tour Guides Under the Demands of Digital Cultural and Tourism Industry Upgrading” (25myjd29).

ACKNOWLEDGEMENTS

Not applicable.

REFERENCES

- [1] Han Yan, "A Critical Thinking Teaching Model for College English Based on Multimodal Theory," *Journal of Ningbo Institute of Education*, vol. 27, no. 2, pp. 109-113, 2025, doi: 10.13970/j.cnki.nbjxyxb.2025.02.025.
- [2] Yuan Jing, "An Exploration of Strategies for Junior High School English Writing Teaching from the Perspective of Multimodal Theory," *Frontiers of Social Sciences*, vol. 14, no. 2, pp. 279-284 DOI: 10.12677/ass.2025.142123, 2025.
- [3] J. Gao and Y. Hua, "On the English Translation Strategy of Science Fiction from Humboldt's Linguistic Worldview-Taking the English Translation of Three-Body Problem as an Example," *Theory and Practice in Language Studies*, vol. 11, no. 2, pp. 186-190, 2021, doi: 10.17507/tpls.1102.11.
- [4] Z. Anwar F, "Strategies and techniques of translation in translating songs as 21st century curriculum," *Manajemen Pendidikan*, vol. 15, no. 1, pp. 34-43, 2020, doi: 10.23917/jmp.v15i1.10744.
- [5] R. Beiler I and J. Dewilde, "Translation as translingual writing practice in English as an additional language," *The modern language journal*, vol. 104, no. 3, pp. 533-549, 2020, doi: 10.1111/modl.12660.
- [6] E. Vula and G. Muhaxhiri, "Exploring linguistic and cultural barriers in literary translation: an analysis of undergraduate students' strategies and challenges in translating Albanian texts into English," *Sapienza: International Journal of Interdisciplinary Studies*, vol. 5, no. 4, pp. e24067-e24067, 2024, doi: 10.51798/sijis.v5i4.864.
- [7] C. Cui, Y. Zhao, and L. Wang, "Protection and dissemination of Chinese intangible cultural heritage based on digital games," *International Communication of Chinese Culture*, vol. 8, no. 4, pp. 483-491, 2021, doi: 10.1007/s40636-021-00237-x.
- [8] O. Guo-zhang Y A, "The Challenges and Solutions for the Digital Dissemination of Intangible Cultural Heritage in China," *Journal of Xihua University (Philosophy & Social Sciences)*, vol. 42, no. 4, pp. 75-82, 2023, doi: 10.12189/j.issn.1672-8505.2023.04.009.
- [9] C. Zhang, Y. Tian, and Y. Zhang, "Analysis of the digital dissemination path of intangible cultural heritage from the perspective of grounded theory," *Advances in Social Behavior Research*, vol. 15, no. 1, pp. 12-15, 2025, doi: 10.54254/2753-7102/2025.20513.
- [10] L. Seifi and M. Soltanabadi, "Iranian public libraries' capacities in preserving and disseminating intangible cultural heritage," *IFLA journal*, vol. 46, no. 4, pp. 359-368, 2020, doi: 10.1177/0340035219886608.
- [11] J. Spagnolo, L. Gautier, F. Champagne, et al., "Reflecting on knowledge translation strategies from global health research projects in Tunisia and the Republic of Côte d'Ivoire," *International journal of public health*, vol. 65, no. 9, pp. 1559-1570, 2020, doi: 10.1007/s00038-020-01502-3.
- [12] B. Yang, A. Gomes Dos Santos, S. Puri, et al., "The industrial design, translation, and development strategies for long-acting peptide delivery," *Expert Opinion on Drug Delivery*, vol. 19, no. 10, pp. 1233-1245, 2022, doi: 10.1080/17425247.2022.2098276.