Proceedings of the 5th International Conference on Literature, Language, and Culture Development
DOI: 10.54254/2753-7064/2026.34340

Analysis of YouTube Self-Media Videos on Chinese Culture and Implications for International Chinese Teaching

Huan Pang

School of Chinese Language and Literature, Shaanxi Normal University, Xi'an , China
panghuan@peihua.edu.cn

Abstract. As an emerging communication medium, self-media videos have become a new carrier for the promotion of Chinese culture. This paper takes 91 popular Chinese culture video texts on YouTube, a leading overseas video platform, as research materials. Through word frequency analysis and TF-IDF analysis, the paper identifies topic words and conducts co-occurrence analysis of topic words to explore the thematic content and emotional attitudes of popular culture videos, so as to provide new ideas for the selection and explanation of themes in international Chinese culture teaching. The study finds that content at the level of cultural history and social life receives high attention and shows positive emotional identity. International Chinese culture teaching can select cultural cases around these hot topics, deepen from the superficial explanation of cultural symbols to the in-depth interpretation of cultural core, and expand from a single language form presentation to multi-modal expression, so as to more effectively convey the connotation of Chinese culture.

Keywords: Self-media, Chinese culture, Theme analysis, International Chinese culture teaching

1. Introduction

Culture is a window for others to understand a nation. Hall first proposed the concept of cross-cultural communication and held that a correct understanding of cross-cultural communication is crucial to promoting exchanges and understanding between different cultures [1]. The effect of cross-cultural communication depends not only on the content and form of information, but also on various factors such as the cultural background and cognitive differences between communicators and audiences. Zhou Qingshan believed that due to the subjective encoding of communicators and the contextual differences of receivers, information may be reconstructed with different meanings in the decoding process [2]. Xu Linjia and Liu Jianhua pointed out that the value of transmitted information will be "discounted" due to cultural differences, language barriers and different historical backgrounds [3]. How to reduce this "discount" has become an important issue. Gillmor, D. put forward *We the Media* and argued that technology has endowed people with a low-cost communication tool to spread information all over the world, and regarded the evolution of this media structure as an "evolution" [4]. Zhang Wenjuan and Gong Chengbo proposed that the multi-form integrated media of self-media platforms (such as text, pictures, videos, etc.) provides

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2753-7064/2026.34340

audiences with diversified experiences, and visualization has become an important representation of online social culture at present [5]. This provides a new idea for international Chinese culture teaching. Selecting cultural topics concerned by others on self-media platforms, designing appropriate cultural content and adopting appropriate explanation methods are effective ways for international Chinese culture teaching to reduce the "discount" of cross-cultural communication.

2. Research design

2.1. Research objects

The steps for selecting video text content are as follows: firstly, search with "Chinese culture" as the keyword on the YouTube web page, select the "video" type in the filter conditions, sort by views, and select the top 100 videos by views; secondly, according to Lu Jianming's definition of "Chinese culture" [6], and in line with the cultural items of *Reference Framework for Chinese Culture and National Conditions Teaching in International Chinese Education*, exclude videos that are not in the category of Chinese culture. After screening, 91 videos are finally determined as analysis samples, with a total of 73,945 Chinese characters after transcription.

The 91 video samples come from 45 account subjects, with registration locations covering 11 countries and regions including the United States, India, China, the United Kingdom and Australia. The languages are mainly English and Chinese, and individual videos are equipped with Hindi subtitles. The samples have accumulated about 740 million views, 150 million likes and 55,000 first-level comments, showing strong communication influence and audience attention. The content covers diverse fields such as history, clothing, food, cities, villages, education, folk customs and traditional Chinese medicine, reflecting the diversity of cultural themes in the samples.

2.2. Research methods

This study uses Python to conduct word frequency analysis, term frequency-inverse document frequency (TF-IDF) analysis and co-occurrence analysis on video sample texts. Word frequency analysis can identify high-frequency words in texts. High-frequency words are related to the main topics of texts. Through word frequency analysis, the thematic content of video texts can be found and the main cultural categories concerned by audiences can be observed. TF-IDF indicates the universal degree of a word in the entire document set. Co-occurrence analysis can reveal the co-occurring words around the topic and further explore the semantic association of the topic. This study selects the top 10 high-frequency words in terms of word frequency and TF-IDF ranking, takes these 10 words as topic words, and conducts co-occurrence analysis with a window of 5. If a topic word appears in a line of text, all other words within the window range of the topic word are regarded as its co-occurring words. If a topic word appears multiple times in a line of text, the co-occurring words within their respective windows are calculated and accumulated each time it appears. When the same co-occurring word co-occurs with a topic word in different text lines, it will be counted separately and merged in the final result. Word order is not distinguished, that is, no matter where a co-occurring word appears in the window, it is regarded as one co-occurrence. The co-occurrence frequency is sorted in descending order, that is, the more times a topic word matches a co-occurring word, the higher the ranking, ensuring that the most representative co-occurrence relationships are displayed first.

DOI: 10.54254/2753-7064/2026.34340

3. Analysis of high-frequency topic words, their co-occurrence and emotions in video texts

3.1. Analysis of high-frequency topic words in video texts

After word segmentation of video texts, removing stop words, conducting word frequency and TF-IDF analysis, and deleting predicate words such as "say" and "do" that cannot display thematic content, the top 10 topic words are obtained. It can be inferred that these 10 words play the role of topic identification in the texts and represent the core themes of the video text content.

Table 1. Topic words of self-media video sample text content

Group Number	Word	Frequency	TF-IDF
1	China	406	0.7209366416899978
2	Work	109	0.19355195552761023
3	Culture	85	0.15093501119125569
4	Ancient	64	0.11364518489694546
5	Traditional	55	0.0976638307708125
6	History	50	0.08878530070073865
7	Life	49	0.08700959468672387
8	House	44	0.07813106461665001
9	City	44	0.07813106461665001
10	Village	41	0.07280394657460569

Classified by cultural components, "China", as a highly complex concept, can be classified into different cultural categories: "China" representing geographical space and cultural material carriers belongs to material culture; "China" involving national governance system, political system and social structure belongs to institutional culture; "China" reflecting national identity, cultural traditions, values and philosophical thoughts belongs to spiritual culture. Words such as "ancient", "traditional" and "history" have cross-level attributes, which can be related to material culture such as ancient artifacts, as well as institutional and spiritual cultural categories such as ancient systems and thoughts. The word "life" covers a wide range, involving not only material culture such as basic necessities of life, but also spiritual culture such as customs and habits. "House", "city" and "village" focus on human living space and settlement forms, which are important components of material culture. "Work" reflects social production, involving labor system, social division of labor, etc., and can be classified as institutional culture.

Classified by time dimension, among the 10 characteristic words, "China" and "culture" are cross-time and space concepts. The remaining words can be divided into two categories: one is related to ancient culture, such as "ancient", "traditional" and "history", reflecting the video content's attention to Chinese history and culture. The other is related to modern life, such as "work", "life", "house", "city" and "village", reflecting the video content's focus on the daily social life of contemporary China.

Through the analysis of topic words in video sample texts, it can be seen that Chinese culture concerned by self-media bloggers and audiences covers both historical categories and content closely related to the social life of contemporary China.

DOI: 10.54254/2753-7064/2026.34340

3.2 Co-occurrence analysis of topic words

Co-occurrence of topic words can clarify the discussion scope of topic words. The author selects the top five co-occurring words with co-occurrence frequency ≥3 for statistics. Repeated co-occurrence groups are counted in the group with the earlier serial number.

Table 2. Co-occurring words and co-occurrence frequency of topic words

Group Number	Theme Word	Co-occurrence Word & Frequency
1	China	Ancient 51, Culture 44, History 22, Shanghai 17, World 15, Traditional 15, Work 15
2	Work	Effort 19, Hours 12, Weekly 10, Life 9, Company 9, Mode 8, Night 8
3	Culture	Impact 9, Martial Arts 7, Tradition 6, Wa Ethnic Group 6, Work 6, History 5, Element 5, Ancient 5, Love 5, Shanghai 4, Food 4, West 4, Shock 4, Experience 4
4	Ancient	History 4, Women 4, Famous 3, Men 3, Style 3, Hanfu 3, Miracle 3, Myth 3, Art 3
5	Traditional	Ancient 9, Dynamic 7, Custom 5, Preserve 4, Cuisine 4, History 4, Wa Ethnic Group 4, Life 4, Courtyard 3, Famous 3, Exploration 3, Festival 3
6	History	Long-standing 10, Years 5, Architecture 4, Story 4, Ancient 4, World Heritage 4, UNESCO 4, Brilliant 4, City Wall 4, Immerse 3, Scenery 3, City 3, Pingyao 3, Myth 3
7	Life	China 11, Way 7, Countryside 6, Happiness 4, Village 4, Joy 3, Exploration 3, Real 3, Like 3, Peaceful 3, Noisy 3
8	City	China 7, Shanghai 7, Peaceful 6, Provide 5, Crowded 4, Culture 3, Miracle 3, Place 3, Developed 3
9	House	China 8, Hotel 6, Visit 6, Beautiful 6, Scenery 5, Village 4, Lovely 4, Farm 4, Interesting 3, Horse 3, Local 3
10	Village	Beautiful 24, China 6, Road 5, Hotel 5, Pretty 4, House 4, Big City 4, Lijiang 4, Exploration 4, Peaceful 3, Snow Mountain 3, Scenery 3

The 10 co-occurrence groups can be divided into "social life" and "cultural history" categories. The former includes topics such as "work", "life", "city", "house" and "village", while the latter includes topics such as "China", "culture", "ancient", "traditional" and "history".

The content discussed around the topic word "China" in Co-occurrence Group 1 mainly focuses on "history", "culture" and "tradition". In particular, the high-frequency co-occurrence of "China" with "ancient" and "culture" indicates that China's long history and rich culture have always been the core content of communication themes. The co-occurrence of "China" with "Shanghai" and "world" reflects the close relationship between China's modernization process and globalization. "Work" involves discussions on individual labor in Chinese society. The co-occurring words of "work" in Co-occurrence Group 2 show that discussions related to work mostly focus on labor intensity (such as "effort", "hours", "weekly") and working environment (such as "company", "mode"). The co-occurrence of "work" with "life" and "night" reflects the close connection between work and daily life, and the issue of work-life balance in modern society has become an important topic. The co-occurrence of "culture" with words such as "impact", "West" and "shock" in Co-occurrence Group 3 indicates that discussions on cultural communication in self-media video themes often adopt a comparative perspective, rather than one-way cultural output. Words such as "martial arts" and "food" indicate that such cultural symbols are more popular. The co-occurrence of words such as "work" and "Shanghai" shows that self-media also involves the context of modern society when discussing culture. The co-occurrence analysis of "ancient" in Co-occurrence Group 4 shows that history, gender and clothing are important discussion points. The co-occurrence of words

DOI: 10.54254/2753-7064/2026.34340

such as "miracle" and "myth" indicates that self-media tends to use legendary narration and positive praise when spreading ancient culture. Words such as "style" and "art" indicate that self-media tends to spread ancient culture from a visual and artistic perspective. The contrast between "ancient" and "dynamic" in the co-occurring words of "traditional" in Co-occurrence Group 5 shows that traditional culture is rooted in China's historical soil and still has vitality. Co-occurring words such as "custom", "cuisine" and "Wa Ethnic Group" indicate that diverse local traditional cultures and customs are also hot topics for cultural communication. Among the co-occurring words of "history" in Group 6, words such as "long-standing", "ancient" and "brilliant" highlight the positive evaluation of self-media videos on the theme of Chinese history and culture. "World Heritage" and "UNESCO" indicate the international recognition of Chinese history and culture. Other words such as "architecture", "city wall" and "Pingyao" highlight the specific cultural symbols concerned by self-media bloggers and audiences under the historical category. The co-occurring words of "life" in Group 7 focus on the description of living conditions and space. "Countryside" and "village" indicate that self-media videos pay high attention to rural life. Words such as "happiness", "joy", "exploration", "like", "real" and "peaceful" indicate that discussions about China's living conditions on self-media generally have positive and optimistic emotional colors. The co-occurrence of "city" in Group 8 shows that cities and spatial characteristics are the focus of bloggers. The co-occurrence of "city" with words such as "peaceful", "crowded", "miracle", "developed" and "place" shows the multiple faces of cities. The high-frequency co-occurrence of "Shanghai" shows its status as a representative city in global cultural exchanges. Group 9 reflects the connection between houses and living space and living experience. The co-occurring words of Group 10 focus on the idealization of rural life and the display of natural beauty, showing the positive and romantic expression of rural life in the communication of self-media videos, and conveying people's yearning for a peaceful, natural and harmonious life.

3.3. Emotional tendency of topic words and co-occurring words

The emotional expression of self-media content directly acts on the audience's cognitive construction and emotional identity. Compared with the communication mode of official media which is rigorous, standardized and focuses on grand narrative, self-media platforms weaken the access threshold and expression constraints of content release, endowing creators with strong creative autonomy and content processing space. Creators can conduct life-oriented, interesting and personalized secondary creation of Chinese cultural symbols, folk customs, social daily life and other content combined with their own life experience and personalized perspective, break away from the traditional grand narrative framework, adopt a more down-to-earth, friendly and infectious discourse way, fit the fragmented and lightweight information receiving habits of overseas audiences, and effectively narrow the psychological distance in cross-cultural communication.

In the practice of spreading Chinese culture for overseas audiences, top self-media creators generally tend to build a narrative framework with a positive emotional tone, mostly adopt a life-oriented and warm positive narrative strategy, take the initiative to avoid negative and sharp expression perspectives, strive to shape a vivid, warm and diverse image of Chinese culture, and weaken the cognitive deviation caused by stereotypes. Combined with the analysis of high-frequency topic words and their co-occurrence relationships in video texts conducted in this study, this emotional tendency can be intuitively confirmed. In the topic word network of sample videos, words full of positive emotional colors such as "effort, miracle, dynamic, happiness, joy, like, lovely, interesting, beautiful, pretty" appear many times in the co-occurring words of high-frequency topic

DOI: 10.54254/2753-7064/2026.34340

words, forming a stable positive semantic association network and continuously conveying an optimistic, inclusive and beautiful cultural temperament.

Based on video views as the screening criterion, this study finds through high-frequency topic word extraction and co-occurrence network visualization analysis that self-media content with international communication influence generally presents a stable and distinct positive emotional tone. This positive narrative not only helps to shape a good impression of Chinese culture, but also makes it easier for audiences to accept and identify with the spread cultural content.

4. Enlightenment of hot topics in self-media videos to international chinese culture teaching

Liu Dazheng proposed that the characteristics of self-media provide more possibilities for language presentation and external communication. With the popularization of mobile terminals among people around the world, self-media plays an increasingly important role in international Chinese teaching [7]. The hot topics of self-media videos can provide a reference for the content selection of international Chinese teaching, and the characteristics of self-media itself can also bring enlightenment to teaching methods.

4.1. Selection of cultural content cases

The video samples selected in this paper have strong communicability, the cultural themes presented in their text content have hot spot attributes, and the thematic content obtained through the analysis of high-frequency words and their co-occurrence is consistent with the cultural items in *Reference Framework for Chinese Culture and National Conditions Teaching in International Chinese Education*. The thematic content of "cultural history" and "social life" corresponds to the cultural items in "traditional culture" and "social life" in the *Reference Framework*. These contents are very suitable for expansion into cases in cultural teaching, which provides an effective idea for teachers to select cultural content, helping teachers arouse students' interest through cultural phenomena and hot topics, and ensuring that the content is close to the times. Around the analysis of high-frequency words and their co-occurrence, specific cultural content can be selected. For example, in historical and cultural categories, around key words such as "history, ancient, traditional", explain Chinese history, Confucian and Taoist culture, classical literature, traditional festivals, etc., to help learners understand the cultural background of Chinese. In social life categories, combined with the development of modern Chinese society, explain Chinese lifestyle, housing, work, urbanization, etc. Discussions can be carried out around topic words such as "life, city, village, house, work" to explore practical issues such as the development of modern Chinese society, urban-rural gap, housing policy, workplace culture, making language and culture learning closer to reality.

Cultural items related to "contemporary China" in the *Reference Framework* are rarely involved in self-media videos. Through the analysis of topic words and their co-occurrence, it can be seen that topic words in fields such as politics, population, social security, education, language and writing, literature and art, technology, media, and foreign exchanges do not appear. This phenomenon conveys a message to international Chinese culture teaching: in the communication from the perspective of others, traditional historical culture and social life content account for a large proportion, and there is an obvious deficiency in the cognition and promotion of contemporary China. Therefore, increasing the explanation of contemporary China in teaching, helping Chinese learners understand contemporary China, spread contemporary China, and tell the story of contemporary China well is an important part of international Chinese culture teaching.

DOI: 10.54254/2753-7064/2026.34340

4.2. From symbol cognition to connotation interpretation

International Chinese culture teaching often revolves around concrete cultural symbols, such as traditional clothing, characteristic food, houses and residences, means of transportation, etc., but these symbols themselves do not directly reflect the deep connotation of culture, and the information they are easy to be recognized and accepted is often superficial. The real cultural interpretation cannot stay at the level of symbol recognition alone, but should go deep into the communication methods, cognitive methods, values and other levels behind the symbols. International Chinese culture teaching needs to conduct in-depth cultural interpretation through symbols, so that learners can not only recognize these cultural symbols and generate interest in cultural interpretation, but also understand the cultural logic and meaning they carry. For example, under the theme of "houses and residences", we can not only introduce traditional courtyard houses, Hakka round houses and other buildings, but also discuss living concepts, housing and the changes of social structure, and guide learners to think about the deep reason why "houses" play an important role in contemporary Chinese society. The reason is that houses are not only living space, but also a microcosm of family concepts, social mobility and changes in economic structure. For another example, under the theme of "traditional customs", we can discuss social order and values, not only introduce customs such as Spring Festival greetings and Chinese weddings, but also further explore the social structure and cultural concepts behind these customs. For example, elders giving lucky money to younger generations during the Spring Festival is not only a festival custom, but also reflects family concepts. The three letters and six rituals in weddings emphasize the social contract between families. The deep reason why traditional customs still have influence in modern society is that customs are not only ritual cultural expressions, but also a reflection of social relations, ethics and identity. It should be noted that this in-depth teaching is more suitable for adult learners at the intermediate and advanced levels.

4.3. From language explanation to multi-modal expression

As an important carrier of modern communication, self-media videos are a force that cannot be ignored in international Chinese education, and their multi-modal communication characteristics are worthy of teaching reference. With this expression form, cultural content can be presented in a more vivid and distinct way. Berlo, D. K. emphasized that the understanding and acceptance of messages are based on the audience's cognitive and decoding processes. Messages are not fixed, but depend on the audience's background, experience and understanding [8]. Multi-modal expression can better help learners understand cultural content. The cultural explanation in international Chinese classrooms needs to shift from a language-led model to situational and visual multi-modal expression integrating images, texts, audio-visual and other elements. International Chinese culture teaching can create a Chinese cultural image of "trustworthy, lovable and respectable", and shape digital characters with distinct characteristics. For example, explain *Along the River During the Qingming Festival* from Zhang Zeduan's perspective, and analyze traditional residential culture from an architect's perspective. For another example, use animation, virtual anchors or youthful expression to lead students to explore Chinese culture. This expression not only makes cultural transmission and interaction more attractive, provides learners with an intuitive learning experience, but also guides them to deeply understand the multi-dimensional content of Chinese society, history and culture, so that learners can understand and identify with Chinese culture in experience.

DOI: 10.54254/2753-7064/2026.34340

5. Conclusion

Taking 91 hot videos on Chinese culture on YouTube as the research object, this paper counts the core topic words through text word frequency analysis and TF-IDF analysis, and explores their co-occurrence characteristics and overall emotional attitude tendency. The research results have certain enlightenment significance for the screening of cultural points and the selection of cultural teaching methods in international Chinese teaching. International Chinese culture teaching is not only the teaching of cultural knowledge, but also an important channel to show China to the world and enhance the international community's understanding of China. By learning from the hot topic words and their co-occurrence rules of self-media videos, and exploring the cultural content concerned and the emotional attitudes of overseas audiences towards hot culture, we can provide a reference for the content design and teaching method optimization of international Chinese culture teaching, and further help to tell Chinese stories well and spread Chinese voice well.

References

[1] Hall, E. T. (2010). The Silent Language. Translated by He Daokuan. Peking University Press.
[2] Zhou Qingshan. (2024). Basic Course of Communication Studies. Peking University Press.
[3] Xu, L. J., & Liu, J. H. (2014). The value evolution in cross-cultural communication: cultural discount and cultural value-added. Chinese Publishing Industry, (08), 8-12.
[4] Gillmor, D. (2004). We the Media: Grassroots Journalism by the People, for the People. O'Reilly Media.
[5] Zhang, W. J., & Gong, C. B. (2020). "Digital faces" in the context of online visual social networking. Contemporary Communication, (03), 98-101.
[6] Lu, J. M. (2015). International Chinese language education and international communication of Chinese culture. Journal of Tongji University (Social Science Edition), 26(02), 79-84.
[7] Liu, D. Z. (2024). A study on strategies of self-media assisting international Chinese teaching. Language and Translation, (01), 72-75.
[8] Berlo, D. K. (1960). The Process of Communication: An Introduction to Theory and Practice. Holt, Rinehart, and Winston.