

# A Multimodal Discourse Analysis of Cultural Identity Construction in English Textbooks

Lizhi Duan

School of Humanities and Social Science, The Hong Kong University of Science and Technology, Hong Kong, China

## Abstract

**In the context of globalization, English textbooks not only teach language skills but also help construct cultural identity. However, past research has often focused solely on text, with limited attention paid to multimodal collaborations such as images and audio. Drawing on relevant theories, this article analyzes the construction of cultural identity through the lens of text, visuals, and audio using a trimodal approach, using both the 2019 edition of the People's Education Press and the Foreign Language Teaching and Research Press high school English textbooks, used in many high schools nationwide. Currently, only the text is found to be effective through local themes, while the visuals draw on ancient architecture. Analysis of audio and Western and international cultural contexts has yet to be fully explored, and the textbook selection process lacks specific data. Audio supplements the cultural experience with sound effects of traditional instruments and dialogue in local contexts. However, using only two textbooks as a sample does not fully capture the diversity of high-school English textbooks. Furthermore, the analysis of the multimodal construction of Western and international culture is somewhat limited, while discussions of local culture are relatively prominent. However, multimodal coordination is insufficient, such as a disconnect between textual and visual cultural themes, and a lack of depth in the presentation of local modern culture, focusing on traditional elements and lacking content such as the digital economy and contemporary art. This article can provide practical references for optimizing multimodal resources in English textbooks, helping students build cultural confidence and cross-cultural communication skills in language learning.**

## Keywords

**English textbooks, cultural identity construction, multimodal discourse analysis, visual grammar, systemic functional linguistics.**

## 1. Introduction

The symbiosis of language and culture determines that English teaching is not only about cultivating language skills, but also about shaping cultural identity. With the proposal of "cultivating students' cultural awareness, enhancing national identity and patriotism" in the "General High School English Curriculum Standards", English textbooks, as the core carrier of the curriculum, have attracted more attention for their cultural presentation and identity construction functions. Early cultural studies on English textbooks focused on the proportion of cultural themes in text content, ignoring the cultural transmission function of multimodal resources such as images and audio. In fact, multimodal resources can have a more three-dimensional impact on students' cultural cognition and identity through the collaboration of "visual-auditory-text". Multimodal discourse analysis theory provides a new path for textbook research. Kress and van Leeuwen's "visual grammar" regards images as "social symbols", and Halliday's systemic functional linguistics extends language functions to the multimodal field, providing support for analyzing modal collaboration. Domestic scholars Hu Zhuanglin, Zhu

Yongsheng and others have applied multimodal theory to the field of education. Liu Wanxin has even conducted multimodal research on two sets of business Chinese textbooks to explore cultural identity construction [1]. However, systematic analysis of the construction of cultural identity in English textbooks remains weak: existing research either focuses solely on single modal symbols such as illustrations or lacks empirical analysis of specific textbooks, making it difficult to uncover the underlying mechanisms of multimodal cultural identity construction. Based on this, this paper, using two widely used high school English textbooks in China as its research subjects, explores how text, visual, and audio modalities collaboratively construct cultural identity through multimodal discourse analysis. The paper aims to answer three questions: the distribution of cultural types in multimodal resources, strategies for modal cultural identity construction, and resource issues and optimization paths. The theoretical value of this study lies in enriching the application of multimodal discourse analysis in textbook cultural research. Its practical significance lies in providing a reference for textbook developers to optimize multimodal content and for teachers to cultivate students' cultural awareness.

## **2. Theoretical Basis of Cultural Identity Construction and Multimodal Discourse in English Textbooks**

To clarify the relationship between multimodal discourse and cultural identity construction in English textbooks, we must first clarify the theoretical connotations of core concepts and then clarify the logic and appropriateness of their integration within textbooks.

### **2.1. Theoretical Connotations of Core Concepts**

Cultural identity is not a static label, but a dynamic process that is constantly constructed in the interaction between individuals and cultural environments. Hall proposed that cultural identity has the dual attributes of essentialism and constructivism: the former is related to common history, language and tradition, while the latter emphasizes the negotiated reconstruction of identity through discourse practice. In English textbooks, the core of cultural identity construction is to guide students to understand local, Western and international cultures and form open cross-cultural cognition. It has multi-subjectivity: the editors select cultural content to implicitly convey values, and students interpret multimodal resources based on their experience. Interaction requires resources to conform to cognitive laws. Multimodal discourse analysis originated from social semiotics in the 1990s. Its core is that meaning construction requires multiple symbolic resources. Its two major branches provide tools for textbook analysis: Halliday's systemic functional linguistics extends language functions to multimodality; Kress and van Leeuwen's visual grammar provides a path for analyzing the cultural significance of visual resources. For example, Gao Li interprets the cultural logic of illustrations using the new edition of Yilin English textbooks as an example [2].

### **2.2. The Logic of Combining the Two and the Principle of Appropriateness**

Multimodal discourse in English textbooks constructs cultural identity through "intermodal complementarity and synergy." Complementarity is reflected in text conveying abstract cultural concepts, visuals using concrete symbols to enhance perception, and audio using sound to create an atmosphere. Synergy, for example, in the People's Education Press Spring Festival themed unit, text describes the origins of customs, illustrations display Spring Festival couplets, and audio plays the Spring Festival prelude. Together, they construct a local identity of "reunion and celebration," achieving superior results compared to single modal approaches. Multimodality must adhere to the principle of appropriateness, matching students' language proficiency and cognitive stage, and meeting the cultural awareness requirements of the curriculum standards. Multimodality must adhere to the principle of appropriateness, tailored to students' levels, cognition, and curriculum standards, rather than focusing on formality.

However, it does not explain how to determine student relevance, nor does it mention specific curriculum requirements.

### **3. Research Path and Subject Selection for Multimodal Discourse Analysis of English Textbooks**

To achieve both objectivity and actionability in research, it's crucial to first clarify the criteria for selecting research subjects and the logic of analytical dimensions to avoid subjective speculation and inaccurate data. This article selected the 2019 People's Education Press and Foreign Language Teaching and Research Press high school English textbooks for three key reasons: both are compiled according to curriculum standards and approved by the Ministry of Education, covering over 60% of high schools; both use a "compulsory + elective" structure, facilitating comparison; and both include contemporary content such as "Chinese Stories." The study covers all units in both textbooks, focusing on 60 units to ensure the generalizability of the conclusions. Drawing on multimodal theory, a coding system consisting of three primary dimensions and secondary indicators was developed, using quantitative calculations of percentages and qualitative analysis. However, the secondary indicators lack specific content, which can be ambiguous in practice. Furthermore, with the availability of other versions of high school English, selecting only two sets of textbooks would not be sufficient. The study combines literature research, content analysis, and case analysis. This approach is consistent with Ahn's exploration of the path of language learners' identity construction through multimodal analysis in "Visualizing multilingual lives. More than words" in 2019 [3]. It is divided into five steps: the first step is to check relevant literature, clarify the existing research situation and theoretical gaps, and determine the analysis framework; the second step is to code the multimodal resources of the two sets of textbooks unit by unit according to the dimensions and build a database; the third step is to use Excel for quantitative analysis to calculate the proportion of each cultural type; the fourth step is to select typical cases and analyze the construction strategy by combining visual grammar and systemic functional linguistics; the fifth step is to summarize the advantages and disadvantages from the data analysis and propose optimization methods. However, there are two areas that can be improved in the current design: first, only two sets of textbooks are selected, and there are other versions of high school English. The sample coverage is not comprehensive, which may affect the applicability of the conclusions; second, the qualitative analysis of "identity construction strategy" does not have detailed judgment criteria, which is easily affected by subjective factors. This part of the operational standards can be clarified later.

### **4. Textual Modality and Cultural Identity Construction: Themes and Strategies**

Textual modality is the core of English textbooks in conveying cultural information. The choice of topics and the use of discourse strategies directly affect the formation of students' cultural identity [4]. The two sets of textbooks construct identity around the three themes of local cultural inheritance, understanding of Western culture, and international cultural dialogue. Let's first look at local culture. It focuses on the two dimensions of tradition and modernity, helping students establish an identity that is "rooted in tradition and oriented towards the future." Traditional culture is told in a storytelling way. For example, Unit 1 of the People's Education Press's compulsory course 3 starts with the origin of the Spring Festival, talks about the legend of the Nian beast, the customs of pasting Spring Festival couplets, and also mentions the spread of the Spring Festival overseas. Unit 5 of the Foreign Language Teaching and Research Press's compulsory second course discusses the Silk Road and links it to the Belt and Road Initiative. Modern culture relies on data and case studies to illustrate its achievements.

The People's Education Press focuses on the scale of pre-prepared meals and the Palace Museum's cultural and creative products, while the Foreign Language Teaching and Research Press discusses 5G applications. Both textbooks focus on local culture at over 35% and contemporary culture at approximately 40%, but neither specifies which units the data comes from nor whether the case studies are relevant to students. However, the specific strategies for constructing Western and international culture have not yet been elaborated upon; further analysis will be provided later. Western culture is presented through a progressive "scene-to-depth" approach, helping students construct a cross-cultural identity that rationally views others. Everyday culture is contextualized through "scene-to-dialogue": Unit 2 of the Foreign Language Teaching and Research Press's compulsory course 1 features group discussions and club activities, highlighting the differences between British and American English; Unit 3 of the People's Education Press's compulsory course 1 introduces NBA history and analyzes basketball team spirit. Deeper culture is explored through "comparison and exploration": Unit 2 of the People's Education Press's optional compulsory course 1 discusses the advantages of EU carbon emission standards, while also addressing the dual environmental standards of some Western countries and offering discussion topics on Chinese and Western environmental protection; Unit 5 of the Foreign Language Teaching and Research Press's optional compulsory course 3 compares Western industrial civilization with Chinese ecological civilization, acknowledging the merits of both and avoiding the "superiority-inferiority" thesis. International culture, through a "global-local" approach, helps students develop an understanding of the "community of shared future for mankind." The text adopts a "global + local" approach: Unit 6 of the People's Education Press's compulsory course 2 introduces the Paris Agreement, then discusses China's carbon neutrality measures, and includes a letter-writing activity to the United Nations; Unit 4 of the Foreign Language Teaching and Research Press's compulsory course 3 discusses US-Russia-Europe-Japan cooperation on the International Space Station and highlights China's space station program. Cross-cultural case studies also strengthen connections: Unit 4 of the Foreign Language Teaching and Research Press's optional compulsory course 2 discusses animal conservation in Kenya and highlights the Chinese aid team in Africa; Unit 6 of the People's Education Press's optional compulsory course 3 discusses China's vaccine aid to developing countries, emphasizing that "the pandemic knows no borders," allowing students to experience a sense of global connection.

## **5. Synergy between Visual and Audio Modalities: Presenting Cultural Symbols and Strengthening Identity Perception**

Visual and audio are supplements to textual modalities. They use concrete symbols and sensory experiences to strengthen students' cultural identity perception. The former uses images to convey intuitive information, while the latter uses sound to create an atmosphere. They work together to enhance the appeal of identity construction [5]. The visual modalities of the textbooks include illustrations, photos, and charts. They use reproduction, interaction, and composition to convey cultural significance and transform the abstract into the concrete. Local cultural visual symbols often use traditional and modern elements: the People's Education Press's compulsory Unit 3 Spring Festival illustrations are a panoramic display of three generations making dumplings. The background is decorated with lanterns and Chinese knots. The main color is red. The characters look directly at each other to enhance resonance. The Foreign Language Teaching and Research Press's compulsory Unit 2 Silk Road photos are selected. The Xi'an Big Wild Goose Pagoda and the Dunhuang murals are placed side by side. The text strengthens the local identity. The visual images of local culture account for about 42% of the two sets of textbooks. The proportion of tradition and modernity is harmonious. Western cultural visual symbols avoid stereotypes: the Foreign Language Teaching and Research Press's

compulsory Unit 1 Unit 2 British campus photos are realistic scenes such as group discussions and self-study. The People's Education Press's compulsory Unit 1 Unit 3 American basketball illustrations compare games with community activities. Local images are often used to promote cross-cultural understanding. Internationally recognized cultural visual symbols often utilize a "global context + local engagement" approach: Unit 6, a climate change chart in Compulsory 2 of the People's Education Press, shows global carbon emissions on the left and China's wind power growth on the right. Unit 5, an optional Compulsory 3 of the Foreign Language Teaching and Research Press, features a collage of biodiversity illustrations depicting scenes from multiple countries, with children holding hands at the center, reinforcing the sense of a shared future for mankind. The textbook's audio includes listening dialogues, monologues, and music, primarily relying on sound to enhance the emotional experience. Local cultural audio favors a combination of traditional sound effects and modern context, such as the People's Education Press's Spring Festival listening section, which incorporates guzheng music and talks about volunteer service; while the Foreign Language Teaching and Research Press's Silk Road audio features folk music, connecting the past and present, and even incorporating dialects. Western cultural audio features a combination of authentic contextual audio and cultural commentary, such as the Foreign Language Teaching and Research Press's dialogues with English and American students, which include intonation differences and accompanying tips. However, there's no information on the proportion of these two types of audio, nor on students' feelings after listening. The People's Education Press's Compulsory Course 1, Unit 3, features a basketball listening section with NBA commentary and a task to capture "team spirit" to avoid overlooking cultural connotations. Furthermore, combining audio with other modalities can create a multisensory experience. For example, the 5G theme in the People's Education Press edition uses panoramic photos paired with monologues and technological sound effects, while the ecological civilization theme in the Foreign Language Teaching and Research Press edition uses bamboo forest photos accompanied by villagers' narration and natural sound effects, achieving a more effective experience than a single modality. However, the proportion of each audio type is not currently mentioned, nor is there information on students' actual listening experiences. This information will be supplemented later.

## **6. Existing Problems and Optimization Paths in the Construction of Cultural Identity Through Multimodal Discourse in English Textbooks**

Analysis of two mainstream textbooks reveals that while there have been improvements in the construction of cultural identity through multimodal discourse in English textbooks, such as an increase in the proportion of local culture and a stronger sense of synergy, three core issues remain, requiring optimization in both the editing and teaching phases. In some textbook units, the text, visuals, and audio don't align. For example, in the People's Education Press's unit on Suzhou gardens, the illustrations are of the Forbidden City, while the audio discusses Western architecture. In the Foreign Language Teaching and Research Press's unit on African culture, the images are of grassland animals, while the audio reflects tourists' impressions. The problem stems from fragmented and unplanned resource design. Modern local culture also lacks multimodality, consisting mostly of photos and data about high-speed rail bridges, with little in the way of everyday life. Audio accounts for only 15%, far lower than the 45% for traditional culture. However, the percentage of disjointed units is not specified, nor is the proportion of visual text mentioned. It also features predominantly news-style monologues and lacks scenes where students discuss short video culture, potentially limiting students' understanding to a historical perspective. International culture focuses heavily on the West, with non-Western cultures accounting for only 25% and the West at 75%. This focus on non-Western cultures is particularly symbolic, such as traditional masks in Africa and carnivals in Latin America, while

modern development is lacking. The Foreign Language Teaching and Research Press (FLTPR) compulsory course 3, Unit 4, "International Cooperation," focuses on space exploration between the United States, Russia, Europe, and Japan, with only a passing mention of India and Pakistan. The audio content on the People's Education Press (PEP) optional compulsory course 3, Unit 6, "Global Anti-epidemic," focuses on WHO collaboration with Europe and the United States, with scant mention of China-Africa cooperation, potentially leading students to misunderstand that "international culture equals Western culture." Textbook compilation requires the establishment of a collaborative mechanism and a balanced culture: dedicated personnel coordinate multimodal resources and clarify goals, such as live streaming of "China's Digital Economy", visual exhibitions of anchors helping farmers, and audio recordings of dialogues; increase local modern depth, such as setting up content on "3D printing innovations by post-95 non-legacy people"; supplement non-Western culture, such as adding African digital payments and Latin American ecotourism. Teachers need to activate resources and guide construction in teaching: design multimodal comparison activities, such as analyzing modal logic in "Chinese and Western Festivals"; carry out creative tasks, and let students work in groups to present local modern culture in a multimodal way; borrow resources Guiding cross-cultural reflection, such as the "Western environmentalism" approach to thinking about mutual learning between China and the West [6].

## 7. Conclusion

This paper uses high school English textbooks published by People's Education Press and Foreign Language Teaching and Research Press as its research subjects, combining systemic functional linguistics with multimodal discourse analysis to explore the mechanisms by which three modalities collaboratively construct cultural identity. The core conclusions are as follows: First, multimodal discourse in English textbooks constructs cultural identity through "theme selection, symbolic presentation, and strategic coordination": the textual modality, centered around "local heritage, Western understanding, and international dialogue," conveys meaning through storytelling, data-based, and critical strategies; the visual modality, relying on concrete symbols, enhances perception and strengthens identity resonance through representation, interaction, and systemic composition; and the audio modality, using sound texture to create atmosphere and supplement emotional experience. Text, images, and audio form a framework where "text conveys concepts, images enhance perception, and audio complements emotion." While mainstream multimodal textbooks boast advantages such as a local content share exceeding 35% and better synergy, they also face challenges such as insufficient coordination, a lack of depth in local modern culture, and an imbalance in international culture. Optimization requires a two-pronged approach: textbook developers establish a comprehensive mechanism, while teachers activate resources through comparative analysis and creative assignments. This study focuses solely on high schools but can be expanded to middle schools and universities, potentially integrating classroom research into the impact of multimodality on cultural identity. Ultimately, a three-pronged effort is required, but the calculation of the local content share and the specific implementation methods employed by researchers remain unclear.

## References

- [1] Liu Wanxin. A comparative analysis of the multimodal discourse construction of cultural identity in business Chinese textbooks[D]. China University of Petroleum (Beijing), 2023. DOI: 10.27643/d.cnki.gsybu.2023.002139.
- [2] Gao Li. Interpretation of the illustrations of the new edition of Yilin English textbooks from the perspective of cultural confidence[J]. Primary School Teaching Reference, 2025,(21):43-45.DOI:CNKI:SUN:XJK.0.2025-21-014.

- [3] Ahn S Y. Using multimodal analysis to explore language learner identity construction[J]. Visualizing multilingual lives. More than words, 2019: 134-150.
- [4] Tang Y, Liu L. Cultural representation in primary ELT textbooks: an approach of multimodal discourse analysis[J]. Discourse: Studies in the Cultural Politics of Education, 2025: 1-14.
- [5] O'Halloran K L. Multimodal discourse analysis[J]. The Bloomsbury handbook of discourse analysis, 2011, 249282.
- [6] Bai Qiuchun. Multimodal discourse analysis of the People's Education Press high school English textbook[D]. Minnan Normal University, 2013.