Educational and Humanities
ISSN Print: 3104-4352
ISSN Online: 3104-4360
CD ISSN: 3104-4379
https://doi.org/10.63313/EH.9093

# Cultivation Strategies for Cultural Awareness in High School English from a Multimodal Perspective

**Jingwen Dong**
China West Normal University, Nanchong 637009, China
Email: 1052323593@qq.com

**How to cite this paper:** Dong, J. (2026). Cultivation strategies for cultural awareness in high school English from a multimodal perspective. Educational and Humanities, 4(1), 45–52. ISSN Print: 3104-4352, ISSN Online: 3104-4360.
https://doi.org/10.63313/EH.9093
**Published: 2026-07-06**

Copyright © 2026 by author(s) and Erytis Publishing Limited.
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
http://creativecommons.org/licenses/by/4.0/
CC BY 4.0

## Abstract

T*he National English Curriculum Standards for Senior High Schools (2017 Edition, 2020 Revised)* identifies "cultural awareness" as one of the four dimensions of core competencies in English language teaching, yet its cultivation is hindered by persistent challenges, including one-sided presentation of cultural knowledge, superficial cultural understanding, and insufficient critical thinking skills. To address these challenges, this paper draws on multimodal theory to examine its compatibility with cultural awareness cultivation, proposing three concrete teaching strategies: utilizing film and video resources to reconstruct cultural contexts, interpreting visual metaphors in images, and employing music and body movements to enhance cultural perception. This paper demonstrates that these strategies effectively guide students from rote memorization toward deeper cultural comprehension, affective engagement, and authentic practice by leveraging the synergistic interplay of visual, auditory, and kinesthetic modes. It concludes that multimodal theory not only offers a viable remedy to the limitations of traditional text-based instruction but also provides a practical framework for designing culturally immersive and cognitively engaging classroom experiences, thereby facilitating the effective implementation of cultural awareness cultivation in senior high school English classrooms.

## Keywords

Multimodal theory; High school English; Cultural awareness; Teaching strategies

## 1. Introduction

*The English Curriculum Standards for Senior High School (2017 Edition, Revised in 2020)* (hereinafter referred to as the Standards) establish cultural awareness as one

of the core competencies in English language teaching. This competency encompasses students' understanding of both Chinese and foreign cultures, identification with valued cultural traditions, and the cross-cultural cognition, attitudes, and behavioral orientations they display in the context of globalization. The Standards explicitly require that cultural instruction should not be confined to the accumulation of factual knowledge; rather, it must prompt students to grasp the deeper cultural significance and to internalize cultural values into their personal beliefs and conduct [1]. This also implies that the cultivation of cultural awareness must move beyond rote memorization toward deeper comprehension, affective engagement, and authentic practice.

In actual senior high school English classrooms, however, persistent challenges remain, including one-sided presentation of cultural knowledge, superficial cultural understanding, and insufficient critical thinking skills. The root cause lies in an overreliance on written texts in classroom instruction, which fails to leverage the educational potential of multimodal semiotic resources—such as images, audio, and video—to support the development of higher-order cultural literacy.

Given that human communication and meaning-making are inherently multimodal, and that cultural transmission and inheritance likewise depend on the coordinated use of multiple semiotic modes, multimodal theory offers a valuable framework. It clarifies the distinctive value and complementary relationship of different modes—language, image, sound, and movement—in meaning construction, thereby offering a viable remedy to the limitations of conventional pedagogical models. In light of the educational imperatives set by the Standards and the current realities of classroom practice, this paper draws on multimodal theory to examine its compatibility with the goal of fostering cultural awareness in high school English teaching. It then proposes practical pathways that are grounded in teaching contexts, with the ultimate aim of effectively implementing the cultivation of cultural awareness in English classrooms.

## 2. The Formation, Development, and Principles of Multimodal Theory

Mode refers to the channels and media of communication, encompassing semiotic systems such as language, technology, image, color, and music[2]. In interaction, people exchange information and manage interpersonal relations through multiple channels and modes, which constitutes the process of meaning-making and transmission. This process is not unimodal but inherently multimodal[3]. Multimodal theory systematically investigates how two or more modes work in coordination to produce holistic meaning.

The systematic development of this theory began in the 1990s. The New London Group first introduced the concept of multimodality, noting that communication is achieved through the interplay of multiple semiotic systems—such as image, sound,

and gesture—rather than through language alone[4]. Kress and van Leeuwen further advanced the theory by proposing the "visual grammar" framework, which provided a comprehensive analytical model for multimodal discourse analysis[5]. In China, Hu Zhuanglin demonstrated the practical value of multimodality in academic and educational contexts[6], while Zhang Delu defined multimodal discourse as communication that employs multiple senses, such as auditory, visual, and kinesthetic, as well as various semiotic resources, including language, pictures, sound, and movement[7].

Building on the ongoing contributions of scholars both at home and abroad, multimodal theory has gradually developed a clear set of principles, which can be understood at three levels. First, language is not the sole resource for meaning-making; non-linguistic modes such as image, sound, color, and gesture are not subordinate to linguistic modes. Each possesses its own semiotic functions and rules, and participates in meaning-making in its distinctive way. Second, the realization of meaning depends on synergy and complementarity among different modes. In other words, when multiple modes are simultaneously engaged in communication, their interplay generates richer and more multidimensional meanings than any single mode could achieve alone. For instance, in advertising, the combination of image and text produces more persuasive effects; in films, dialogue paired with background music conveys emotions that language alone cannot fully express. Third, the selection and use of modes are deeply conditioned by cultural environments, as different cultures exhibit considerable differences in how color, space, body language, and other modes are perceived and employed.

## 3. The Compatibility of Multimodal Theory with the Cultivation of Cultural Awareness in Senior High School English

From the perspective of the nature of culture, it can be understood as a complex yet systematic ensemble of meanings. Culture not only serves as a vital vehicle for human communicative activities but also possesses an inherently multimodal character. Taking the traditional Chinese Spring Festival as an example, its spiritual core is preserved and transmitted not only through written texts such as classical documents but also through various concrete and symbolic forms that together construct a complete cultural meaning system. Spring Festival couplets, written in black ink on red paper, carry auspicious wishes for warding off evil and inviting good fortune; the sound of firecrackers, as a typical auditory symbol, embodies the collective cultural memory of bidding farewell to the old and ushering in the new; and the traditional ritual of younger generations bowing to their elders manifests the moral value of filial piety and respect for the aged. Evidently, these visual, auditory, and kinesthetic modes do not function in isolation; rather, through their interplay, they jointly constitute the rich and multi-dimensional spiritual core of the Spring Festival.

Cultural learning itself is not a matter of textual study alone. Beyond acquiring basic linguistic knowledge, learners are expected to develop the ability to identify various cultural symbols and interpret meanings conveyed through multiple modes. In the context of senior high school English teaching, the instructional texts that carry both linguistic and cultural knowledge are largely multimodal in nature. Relevant studies have confirmed that modes such as text, image, and color complement and coordinate with one another, jointly shaping the multimodal features of textbook materials, thereby enabling both students and teachers to comprehend and internalize knowledge from multiple perspectives during the reading process[8]. This implies that the learning and understanding of cultural knowledge require teachers, on the basis of careful analysis of the multimodal features of texts, to guide students in deconstructing multimodal symbols. Meanwhile, as a component of the core competencies in English language teaching, cultural awareness itself exhibits a hierarchical structure. Scholar Liu Daoyi further delineated cultural awareness into five dimensions: cultural knowledge, cultural consciousness, cultural attitudes, cultural character, and cross-cultural communicative competence[9]. The Standards also explicitly state that the cultivation of cultural awareness should enable students to internalize these values into their personal beliefs and conduct[1]. Such a deep-level objective is difficult to achieve through text-based teaching alone; it necessitates moving beyond the confines of purely text-based instruction and leveraging multimodal teaching approaches, so that students can perceive and understand the connotations of different cultures in authentic or simulated cultural contexts.

Specifically, multimodal theory can provide a solid practical foundation for the cultivation of cultural awareness in senior high school English classrooms at three levels: situational reconstruction, connotative interpretation, and affective resonance. First, multimodal input can reconstruct cultural contexts to a high degree. In English film clips, elements such as visual composition, character costumes, and background music collectively convey the values and lifestyles of a given culture. Through activities such as audio-visual appreciation and classroom discussion, students can gain a more intuitive, comprehensive, and multi-faceted cultural experience. Second, the visual metaphors embedded in images and symbols can help students delve into the deeper meanings of culture. Many abstract cultural concepts are often metaphorically expressed through visual modes such as color, spatial arrangement, and shape—for example, green symbolizing nature and environmental protection; symmetrical layouts signifying balance and order; and circular forms representing reunion and harmony. Interpreting such visual metaphors facilitates students' understanding of the underlying logic and spiritual essence of different cultures. Third, modes such as music and body movement can effectively stimulate students' affective resonance with diverse cultures. Cultural learning entails both cognitive understanding and emotional identification; the

incorporation of auditory and kinesthetic modes helps students gradually deepen their cultural perception through active participation.

In light of the above analysis, multimodal theory demonstrates a high degree of compatibility with the cultivation of cultural awareness in senior high school English classrooms, thereby offering a viable pathway for addressing the persistent challenges in traditional teaching practices.

## 4. Cultivation Strategies for Cultural Awareness Based on Multimodal Theory

Drawing on the core principles of multimodal theory, this section proposes specific strategies for cultivating cultural awareness from three dimensions: utilizing film and video resources to reconstruct cultural contexts, interpreting visual metaphors in images, and employing music and body movements to enhance cultural perception.

### 4.1. Utilizing Film and Video Resources to Reconstruct Cultural Contexts

Film and video resources integrate multiple modes, such as visual and auditory, and can vividly reconstruct various cultural scenes, offering students an immersive cultural experience. When selecting such instructional materials, teachers should adhere to three basic principles. First, teachers should ensure cultural authenticity by prioritizing clips that accurately represent the target culture and avoiding stereotypes. Second, they should ensure pedagogical suitability by selecting clips based on instructional objectives and students' language proficiency, with reasonable control over linguistic difficulty and the provision of bilingual subtitles when necessary. Third, they should ensure modal richness by focusing on clips in which different modes work in coordination, thereby providing sufficient material for multimodal analysis.

The teaching procedure using film and video resources can be divided into three steps. The first step is focused perception, in which teachers can guide students to attend to multimodal signs in the clips through a sequence of guiding questions, such as "What are the characteristics of the characters' costumes in this scene?" and "What feelings does the background music evoke?" The second step is collaborative analysis, in which students are guided to examine how different modes work together to convey cultural information. For example, after watching clips related to Western festivals, students can intuitively perceive multiple modal signs—such as warm color tones, cheerful music, and family embraces—and then further analyze how these three elements jointly construct the theme of "reunion and gratitude". The third step is contextual reconstruction, in which students, building on their prior understanding of the cultural theme, engage in role-playing or scenario-based performance activities in groups, producing creative cultural expressions in

simulated contexts and thereby deepening their comprehension of the cultural theme.

### 4.2. Interpreting Visual Metaphors in Images

Images convey deep cultural meanings through visual elements such as color, composition, and shape. Therefore, interpreting images not only cultivates students' visual literacy but also helps them grasp the metaphorical logic of cultural symbols and build a systematic framework of cultural understanding. When selecting images, teachers should also observe three basic principles. First, the image should carry cultural symbolism. For instance, when teaching content related to drama, teachers may choose Peking Opera facial masks as the object of analysis, guiding students to understand that different colors of the masks metaphorically represent the moral character and personality traits of the figures. Second, the modes should be diverse, with preference given to symbols that integrate multiple modes such as text, color, and shape. For example, illustrations comparing Chinese and Western drama should convey their artistic differences through various modes, including stage design, costume patterns, and performance postures. Third, the image should be comprehensible, ensuring that its visual metaphors align with senior high school students' cognitive level and life experience, and avoiding symbols that are overly abstract or too strongly region-specific.

To further deepen students' understanding of visual symbols and promote the development of cultural awareness, teachers may, on the one hand, encourage students to engage in image reconstruction. This involves creating works with visual metaphors around specific cultural concepts and then offering interpretive explanations. For example, students may be guided to use the Taiji diagram to symbolize the harmonious coexistence between humans and nature, enabling them to grasp, through the creative process, the constructive logic between visual symbols and culture. On the other hand, teachers may guide students to conduct cross-cultural comparison activities based on multimodal content, such as comparing the representational forms of traditional Chinese festivals and Western festivals in visual modes, and then creating cultural comparison posters. Such activities not only encourage critical output and cultivate students' thinking skills but also enable them to deepen their understanding of traditional Chinese festivals, strengthen their sense of identification with these festivals, and foster cultural confidence.

### 4.3. Employing Music and Body Movements to Enhance Cultural Perception

Auditory stimulation and physical experience serve as important bridges for connecting with cultural emotions and as key links in cultivating cultural awareness. Musical works nurtured by different regional cultures carry distinctive historical

depth and emotional connotations in their melodies, rhythms, and instrumentation. Similarly, body movements—such as daily social etiquette and festive dances—intuitively reflect the communicative norms and value orientations embedded in specific cultural contexts.

When integrating music into English classroom teaching, teachers should follow two principles. On the one hand, they should prioritize musical works with cultural representativeness, using classical pieces to establish connections between musical art and regional cultures, thereby helping students gradually develop cultural and emotional identification. On the other hand, they should emphasize modal integration by combining music with other modes, such as movement, to fully engage students' multiple senses.

Conducting cultural teaching through body movements also holds substantial educational value. Teachers may begin by designing relatively simple activities of movement imitation and cultural comparison, such as guiding students to imitate common greeting etiquette in different countries. Through such embodied practice, students can experience how the abstract concept of "respect" manifests in various cultural contexts. In addition, teachers may design more elaborate situational performance activities. Unlike imitation activities, which involve reproductive output, situational performance requires students to apply their acquired cultural knowledge creatively in realistic scenarios. Therefore, such activities are more conducive to cultivating students' creative thinking.

## 5. Conclusion

As a value-oriented component of the core competencies in English language teaching, the cultivation of cultural awareness should move beyond the purely text-based approach and explore more multidimensional and diversified teaching pathways. Multimodal theory systematically elucidates the distinctive value and synergistic mechanism of various semiotic systems—including language, image, sound, and movement—in meaning-making, and in doing so, offers both new theoretical perspectives and practical tools for cultural teaching in senior high school English classrooms. Therefore, in the process of cultivating students' cultural awareness, teachers should move from unimodal instruction to multimodal collaborative teaching. This approach guides students to progress from rote memorization of cultural knowledge toward deep construction of cultural connotations, and from passive reception toward active expression, ultimately facilitating the realization of cultural awareness in practice

## References

[1] Ministry of Education of the People's Republic of China. (2020) English Curriculum Standards for Senior High School (2017 Edition, 2020 Revised). Beijing: People's Education Press.

[2] Zhu, Y.S. (2007) Theoretical Foundations and Research Methods of Multimodal

Discourse Analysis. *Foreign Language Research*, (05), 82-86.

[3] Hu, Z.L. (2006) The Multimodal Construction of Meaning. *Technology Enhanced Foreign Language Education*, (08), 3-12.

[4] New London Group. (1996) A Pedagogy of Multiliteracies: Designing Social Futures. *Harvard Educational Review*, 66(01), 60-92.

[5] Kress, G., & van Leeuwen, T. (2006) Reading Images: The Grammar of Visual Design (2nd ed.). London: Routledge.

[6] Hu, Z.L. (2007) Multimodality in Social Semiotic Studies. *Language Teaching and Linguistic Studies*, (01), 1-8.

[7] Zhang, D.L. (2009) Exploring a Comprehensive Theoretical Framework for Multimodal Discourse Analysis. *Foreign Languages in China*, 6(01), 24-30.

[8] Li, M. (2021) A Multimodal Analysis of Textbook Discourse: A Case Study of Oxford Shanghai Junior High English Textbooks. *Foreign Language Teaching in Schools (Middle School Edition)*, 44(10), 48-53.

[9] Liu, D.Y. (2018) On Cultural Awareness as a Core Competency in English Language Teaching. *Research on Classroom Teaching in Primary and Secondary Schools*, (06), 3-8.