International Journal of Education and Humanities
ISSN: 2770-6702 | Vol. 25, No. 2, 2026

Application and Reflection of Generative Artificial Intelligence in Empowering Multimodal Teaching of International Chinese Education

Xin Jin

Hong Kong Metropolitan University, Hong Kong, China

Abstract: The development of digital education technology has effectively promoted the innovation of the international Chinese language teaching model and improved the problem of the rigidity and lag of traditional teaching. Generative artificial intelligence can generate teaching resources, achieve human-computer interaction, and dynamically adjust teaching content and pace based on the learning situation, effectively compensating for the shortcomings of the traditional offline classroom, such as the single resources, fixed mode, and difficulty in implementing personalized teaching. Currently, overseas Chinese language teaching is still mainly based on paper textbooks and text lectures. The single teaching carrier and teaching form are difficult to adapt to the cognitive characteristics of learners from different countries and cannot meet the differentiated learning needs, which restricts the improvement of the quality of overseas Chinese language teaching. Multimodal teaching integrates various media resources such as pictures, text, audio, and video, breaks through the limitations of pure text teaching, conforms to the cognitive laws of second language acquisition, and is more suitable for overseas Chinese language teaching scenarios. This teaching mode mobilizes the multi-sensory participation of learners in learning, helps them build a complete Chinese knowledge system, and effectively improves the teaching effect. Based on the first-line overseas Chinese language teaching practice and research data, this paper focuses on analyzing the application status of generative artificial intelligence in multimodal Chinese language teaching, summarizing the existing practical problems of current technology application, and proposing optimization countermeasures in combination with real teaching situations, providing a reference for the intelligent upgrade and high-quality development of international Chinese language education.

Keywords: Generative Artificial Intelligence, International Chinese Education, Multimodal Teaching, Teaching Innovation, Artificial Intelligence.

1. Introduction

The deep integration of digital technology with the education industry has led to the gradual advancement of international Chinese education towards intelligence. International Chinese education not only aims to cultivate the Chinese language skills of foreign learners but also aims to disseminate Chinese culture and promote cultural exchanges between China and other countries. The number of people learning Chinese worldwide is constantly increasing. Students from different countries and with different backgrounds have significant differences in their cognitive methods and learning habits. Targeted localized teaching adjustments are the key content of current Chinese language teaching reforms [1]. Most overseas Chinese language classes in China still rely on paper-based textbooks for teaching, with a single teaching form, lacking real language communication scenarios, slow content update speed, and inability to create an immersive language learning environment, which restricts the improvement of students' Chinese application abilities and cross-cultural literacy.

The dual encoding theory and the situational learning theory provide theoretical support for the implementation of multimodal teaching in the classroom. Traditional teaching is mainly based on pure text explanations, while multimodal teaching incorporates various media elements such as pictures, audio, and situational interactions. Through a learning mode that combines multiple senses, it helps students establish a complete Chinese cognitive system. Generative artificial intelligence can cover the entire teaching process, from pre-class preparation, classroom teaching, teacher-student interaction, to teaching evaluation and after-class consolidation, providing technical support. The system can create multimodal teaching materials based on teaching needs, rely on big data analysis to understand students' learning situations and carry out personalized guidance, solving many practical problems in foreign language Chinese teaching. Intelligent technology is adaptable to various teaching scenarios, promoting the upgrading of international Chinese education towards intelligence and precision. The integration of technology and teaching is still in the initial exploration stage. Problems such as poor adaptability of teaching resources, poor classroom integration effects, weak intelligent teaching abilities of teachers, and the absence of supporting systems will all affect teaching quality [2]. This article combines first-line teaching experience and existing research results, sorts out the application scenarios of intelligent technology and existing problems, builds an optimized solution suitable for foreign language Chinese classrooms, and helps the standardized development of intelligent Chinese teaching.

2. The Core Value of Generating Artificial Intelligence Enabling International Chinese Multimodal Teaching

In daily teaching, teachers can utilize generative artificial intelligence, combining the key points and difficulties of the class and the learning differences of students from different

countries, to customize multimodal teaching resources that are suitable for the actual classroom. The rich classroom forms and diverse practice scenarios can make up for the shortcomings of traditional Chinese teaching and steadily improve the quality of foreign language classes [3].

There have always been problems of insufficient contextual resources and complex administrative tasks for teachers in overseas Chinese teaching. These two factors have long limited the improvement of teaching quality. General Chinese textbooks in China focus on systematic theoretical knowledge organization, which has a wide range of applicability but insufficient targeting. They lack teaching content relevant to the life scenarios of overseas students. Teachers need to spend a lot of time creating courseware, organizing exercises, and collecting learning data to compress the time for teaching design, one-on-one tutoring, and teaching research, resulting in a significant reduction in the targeted nature of classroom teaching [4]. Generative artificial intelligence can quickly generate various scenarios of teaching materials, solving the problems of single scenarios, outdated content, and weak country adaptability of traditional textbooks, and further improving the resource system for foreign language teaching.

The assistance of generative artificial intelligence in Chinese teaching mainly lies in optimizing teaching resources, achieving precise teaching, and reducing teachers' burdens. The intelligent system can continuously collect students' learning data in class exercises, after-class training, and stage tests, accurately identify students' weak points in word accumulation, grammar application, oral expression, and writing output, and formulate personalized improvement plans for different levels of learners, implementing differentiated teaching. The AI can also replace teachers to complete material production, homework grading, and learning data statistics, reducing teaching administrative pressure and allowing teachers to focus more on classroom innovation, cultural interpretation, and personalized tutoring, thereby enhancing the practical effect and humanistic value of classroom teaching [5].

3. The Specific Application of Generative Artificial Intelligence in International Chinese Multimodal Teaching

3.1. Intelligently Generate Multimodal Teaching Resources to Enrich the Teaching Carrier System

The implementation effect of multimodal teaching mainly depends on the matching degree between the quality of teaching resources and the learning situation of students. Traditional Chinese teaching mainly relies on teachers to manually create and integrate multimodal teaching resources, resulting in relatively low overall efficiency. The resources produced manually are in a fixed form and have insufficient depth of content, making it difficult to meet the differentiated learning needs of students with different Chinese proficiency levels. Generative artificial intelligence can effectively address the shortcomings of traditional resource construction and comprehensively cover core aspects of Chinese teaching such as speech correction, character cognition, grammar analysis, writing training, and cultural integration. It can generate professional and scenario-based multimodal teaching resources on demand. For the learning difficulties of overseas students in aspects such as character structure, standard pronunciation, and grammar logic, this technology can combine dynamic demonstrations, standard audio, and real scene images to break down knowledge points, presenting abstract language rules intuitively, reducing the difficulty for zero-beginner learners, and helping students quickly solidify their Chinese foundation and establish a complete language knowledge system [6].

Cross-cultural teaching is an important part of international Chinese education, directly affecting the language acquisition and cultural cognition effects of overseas learners. Cultural differences between China and foreign countries can cause students' misunderstandings, while high-quality multimodal resources can effectively eliminate cross-cultural cognitive barriers. With the help of generative artificial intelligence, teachers can quickly create graphic and video materials on themes such as folk culture, regional landscapes, and social life based on the teaching progress and students' proficiency levels, enriching the forms of classroom teaching. Compared to traditional text-based textbooks, multimodal resources combining audio and video can construct concrete learning scenarios, helping students intuitively understand the cultural connotations carried by characters, reducing cognitive misunderstandings caused by cultural differences. During the teaching process, teachers can flexibly adjust the difficulty and presentation form of teaching resources according to individual characteristics such as students' nationality background and Chinese proficiency. The dynamic and adaptable application method can improve the utilization rate of teaching resources, enhance the targeted and effectiveness of international Chinese classroom teaching.

3.2. Build Immersive Interactive Scenarios and Revolutionize the Teaching Model

The language application proficiency in real-life situations is a key indicator for evaluating the effectiveness of second language acquisition and is also the core educational goal of foreign language curriculum construction. Currently, there is a significant bias in the teaching arrangement of foreign language courses both domestically and internationally. Teaching resources are mostly concentrated on the teaching of theoretical knowledge and the training of exam question types, while the practical teaching focused on oral output has been long neglected. The long-term imbalance in the teaching layout has led to the current situation of polarized development of students' language abilities. The ability to answer exam questions is generally up to standard, but the oral communication ability in real-life scenarios is generally lacking. The fixed traditional oral teaching mode is the main cause of this problem. Classroom training mostly relies on preset questions and mechanical repetition of texts, and the training scenarios are highly formulaic, disconnected from the real-life, socialized and natural communication context. Therefore, the language knowledge acquired by students in the classroom is difficult to complete scene migration and cannot flexibly adapt to diversified social scenarios such as foreign-related interactions and workplace communication. The practical value of language teaching cannot be implemented. The application of generative artificial intelligence in teaching provides technical support for solving the inherent problems of traditional oral teaching. Relying on powerful scene simulation capabilities, this technology can replicate various mainstream communication scenarios, and

learners can conduct immersive oral practical exercises at any time. It effectively fills the teaching gap of scarce teaching resources and insufficient practical exercises in traditional teaching [7].

The deep integration of intelligent hardware devices and artificial intelligence technology breaks through the inherent limitations of traditional oral teaching relying on the classroom and class hours, providing a new carrier for the autonomous and regular oral training of second language learners. The intelligent learning platform has real-time voice collection and intelligent analysis functions, which can accurately identify various problems in learners' oral output, such as pronunciation deviations, grammatical errors, sentence structure and order confusion, and expression logic confusion, and provide targeted correction plans and optimization guidance simultaneously, helping learners continuously polish the quality of oral output and gradually improve the standardization and fluency of language expression. Compared with the reserved atmosphere of traditional classroom teacher-student interaction, the human-machine collaborative practice mode has the advantages of low pressure and high freedom, which can effectively eliminate the psychological barriers of students' fear of speaking and refusal to express, and significantly enhance their subjective willingness to participate in oral training. Learners can focus on core sentence patterns in high-frequency scenarios such as daily social interaction and workplace communication for targeted training, gradually improving their personal language system and comprehensively enhancing their second language communication literacy. At the same time, the immersive intelligent teaching mode reconstructs the traditional one-way transmission teaching form, fully activates students' classroom participation vitality, and significantly improves the overall teaching quality and learning efficiency of oral courses [8].

3.3. Carry Out Personalized Emotional Assessment, and Promote the Implementation of Tiered and Precise Teaching

International Chinese learners come from various countries around the world, with different ages, learning backgrounds, and learning habits. Each person's knowledge gaps and learning goals also have significant differences. The unified collective teaching model cannot simultaneously meet the differentiated needs of beginners, international students, and professional Chinese language learners. Personalized teaching has always been the key and difficult point in foreign language Chinese teaching. The big data analysis function of generative artificial intelligence can completely record all the data of students' classroom answers, oral training, homework, and stage assessments. Through intelligent analysis, it can accurately identify individual learning problems and generate detailed learning situation reports, providing reliable data support for precise teaching and one-on-one tutoring [9].

Learning situation analysis data is an important basis for implementing differentiated teaching. The intelligent platform will push corresponding learning resources and training contents based on students' Chinese proficiency. New students with weak foundations can rely on the platform's phonetics, words, basic grammar, and other consolidation content to build a solid language foundation; advanced students with outstanding abilities can expand their knowledge reserves and improve language application skills through literary appreciation, creative writing, industry Chinese, and cross-cultural communication. This teaching model conforms to the differentiated learning laws of second languages and improves the problems of ambiguous teaching stratification and poor remedial and improvement effects in traditional teaching, making classroom teaching more targeted.

4. The Practical Challenges of Generating Artificial Intelligence Enabling Multimodal Teaching

The continuous update of intelligent technologies has provided a new direction for the innovation of multimodal teaching. The integration of technology and teaching is still in the exploratory stage, and there are many shortcomings in practical applications. Problems such as untimely update of training data, limited coverage of language materials, and loopholes in algorithm logic have led to situations where some AI-generated teaching materials have awkward sentences, inaccurate grammar, and one-sided cultural interpretations [10]. Some intelligent scene materials are disconnected from contemporary Chinese social life and are directly used in foreign language teaching, which may cause overseas students to form a one-sided language and cultural perception and fail to meet the rigorous and professional teaching requirements of cross-cultural teaching.

Most teachers' intelligent teaching abilities are relatively weak, and their application of technology only stops at the surface. This is the main factor hindering the deep integration of technology and teaching. Most teachers only master the basic operations of AI material production and have not formed a systematic thinking for intelligent teaching design, unable to deeply combine intelligent resources with course goals, teaching links, and students' learning situations. The application of technology can only simply enrich classroom forms but cannot effectively improve teaching quality. Currently, there is no unified industry standard for AI Chinese teaching in China, and related systems such as resource copyright, data security, and content review are not yet fully developed. This can easily lead to copyright disputes, data leakage, and other security problems, which is not conducive to the long-term standardized development of intelligent Chinese teaching [11].

5. Generative Artificial Intelligence Enables the Optimization Path of Multimodal Teaching

The quality of teaching resources directly determines the professional level and teaching effectiveness of cross-cultural teaching. In the current situation where the quality of AI resources varies and the adaptability to students' learning situations is insufficient, teachers cannot simply apply intelligent materials in their teaching and need to establish a regular resource review and screening mechanism. Before class, they should comprehensively review the multimodal resources such as images, videos, and audio generated by AI from four dimensions: language norms, cultural interpretation, logical structure, and student adaptability, correct the non-standard expressions and one-sided cultural interpretations, eliminate content that is detached from real-life scenarios, and select high-quality resources for classroom use [12].

Local adjustments can effectively improve the quality of foreign-related intelligent teaching. Teachers should combine the key points of the textbook, the difficulties of cross-cultural teaching, and the cognitive characteristics of students from different countries to refine the instructions generated by AI, precisely control the content direction, difficulty level, and presentation form of the resources. Through refined instruction design and multi-level review, they can develop localized multimodal teaching resources that are adapted to the learning situations of different countries, avoid teaching deviations caused by general materials, and make intelligent teaching content more closely match the actual needs of foreign-related classrooms [13].

The professional ability of teachers is the core guarantee for promoting the deep integration of technology and education. Universities and foreign-related teaching institutions need to build a regular teacher training platform, carry out special training on practical operation of AI tools, multi-modal classroom design, student learning data analysis, and cross-cultural intelligent teaching, fill in the gaps in teachers' intelligent teaching capabilities, and cultivate systematic teaching research thinking. This will lay a solid foundation for the integration of technology and education [14].

Teachers need to establish a scientific concept of technology use and clearly define the auxiliary tool attribute of artificial intelligence. Intelligent technology can simplify teaching processes and reduce work pressure, but it cannot replace the cultural guidance and emotional education value of teachers.

Teachers need to break away from the single mindset of material production, integrate intelligent technology into all teaching links such as lesson planning design, classroom interaction, after-class improvement, and learning situation review, and build a complete teaching system that integrates resource production, immersive teaching, and data-driven learning situation analysis. This can maximize the effect of intelligent technology [15].

The cultural background, cognitive habits, and language foundation of students from different countries vary greatly. The general intelligent teaching model cannot be adapted to all students. During teaching, teachers should combine the characteristics of the students' native language, cultural cognition weaknesses， and language learning weak points to customize differentiated multimodal teaching plans, avoid cross-cultural learning misunderstandings, improve teaching professionalism and adaptability, and make intelligent technology transform from surface application to a core driving force for improving teaching quality [16].

Complete institutional norms are the foundation for the stable and long-term development of intelligent teaching. Education authorities should optimize relevant industry standards and regulations based on the intelligent development trend of international Chinese education, clarify the copyright ownership of AI teaching resources, data usage rules, and privacy protection requirements, refine content review standards, and establish a hierarchical review and regular supervision system to prevent various risks such as content deviation, data leakage, and copyright disputes [17].

On the basis of unified industry norms, each foreign-related teaching institution should, based on its own school characteristics and students' learning situations, formulate internal rules for the use of AI classrooms, clearly define the applicable scenarios and boundaries of AI technology, and avoid issues such as technology abuse and formalistic teaching. Through the two-way improvement of industry norms and internal rules, a comprehensive guarantee system can be established to ensure the continuous development of intelligent multimodal teaching.

6. Conclusion

The wide application of generative artificial intelligence has improved the problems of monotonous traditional international Chinese teaching resources, lack of real-scene scenarios, insufficient personalized teaching, and poor cross-cultural adaptability. It has promoted the innovation and upgrading of multimodal teaching models, and enabled international Chinese education to steadily develop in the direction of intelligence, localization, and diversification. Intelligent technology is merely a teaching auxiliary tool. The humanistic guidance, cultural, and emotional education of teachers are irreplaceable. Language ability cultivation, knowledge imparting, and cross-cultural communication have always been the core goals of international Chinese education.

The current practice of integrating technology and education still has many shortcomings, such as low resource compatibility, insufficient intelligent teaching ability of teachers, insufficient depth of technology integration, and incomplete institutional guarantees. These issues directly affect the development quality of intelligent teaching. Educators need to objectively view the advantages and disadvantages of intelligent technology and continuously optimize the intelligent multimodal teaching model through measures such as strictly controlling resource quality, strengthening teacher capabilities, deepening technology-education integration, and improving supporting mechanisms. This will release the educational value of technology. The optimized teaching system can effectively improve the overall quality of international Chinese education, facilitate the overseas dissemination of Chinese culture and cultural exchanges between China and foreign countries, and provide lasting impetus for the innovation of the discipline.

References

[1] van der Zee, T., & Reich, J. (2024). Large language models in language teaching: A scoping review of empirical research. *Language Teaching*, 57(4), 413–435. https://doi.org/10.1017/S0261444823000458

[2] Sockett, G., & Manganelli, M. (2024). Generative AI for oral interaction in second languages: Learner perceptions and task design principles. *ReCALL*, 36(1), 56–74. https://doi.org/10.1017/S095834402300022X

[3] Hung, Y. W., & Lin, C. H. (2022). Multimodal annotation for Chinese character learning: Effects on orthographic awareness. *ReCALL*, 34(2), 186–204. https://doi.org/10.1017/S0958344021000241

[4] Lim, F. V., & Ng, W. K. (2014). Multimodal task-based language learning with digital media: A framework for Chinese as a foreign language. *System*, 45, 78–91. https://doi.org/10.1016/j.system.2014.05.004

[5] Kress, G., & van Leeuwen, T. (2021). Multimodality and social semiotics: Revisiting the theory for digital media. *Multimodal Communication*, 10(2), 131–150. https://doi.org/10.1515/mc-2020-0027

[6] Bezemer, J., & Kress, G. (2016). Multimodal pedagogy in digital learning environments: A social semiotic perspective.

Learning, Media and Technology, 41(4), 411–425. https://doi.org/10.1080/17439884.2015.1108929

[7] Rokkones, R. K., & Giannakos, M. (2025). Toward hybrid teaching intelligence: Investigating the potential of teacher–AI collaboration using large language models. *Behaviour & Information Technology*, 44(14), 2987–3004. https://doi.org/10.1080/0144929X.2024.2412453

[8] Estevez-Ayres, I., Callejo, P., Hombrados-Herrera, M. A., et al. (2024). Evaluation of LLM tools for feedback generation in programming and language learning contexts. *International Journal of Artificial Intelligence in Education*, 34(3), 487–514. https://doi.org/10.1007/s40593-024-00372-1

[9] Alzahrani, S., & Oyelere, B. (2024). A systematic review of AI-assisted vocabulary learning in second language education. *Education and Information Technologies*, 29(12), 15789–15814. https://doi.org/10.1007/s10639-024-12467-0

[10] De Costa, P. I., & Li, G. (2023). Ethical considerations of generative AI in multilingual language education. *Journal of Multilingual and Multicultural Development*, 44(9), 789–802. https://doi.org/10.1080/01434632.2023.2224301

[11] Jiang, L., & Lai, C. (2024). Digital multimodal composing in second language learning: A systematic review. *Language Teaching Research*, 28(5), 1102–1126. https://doi.org/10.1177/13621688231194786

[12] Lan, Y. J. (2024). AI literacy for language teachers: A framework for responsible integration of generative AI in second language classrooms. *Educational Technology & Society*, 27(3), 112–126. https://doi.org/10.14789/jat.27.3.9

[13] Liu, M., Zhang, L. J., & Biebricher, C. (2024). Investigating students' cognitive processes in generative AI-assisted digital multimodal composing and traditional writing. *Computers & Education*, 211, 104977. https://doi.org/10.1016/j.compedu.2024.104977

[14] Eguchi, M. (2024). Designing LLM-mediated role-play tasks for L2 pragmatic competence. *Language Learning & Technology*, 28(2), 108–129. https://doi.org/10.1016/j.lilt.2024.02.006

[15] Yang, Y. (2025). AI-supported L2 vocabulary acquisition–a systematic review from 2015 to 2023. *Education and Information Technologies*, 30, 18241–18267. https://doi.org/10.1007/s10639-025-13412-3

[16] Pellerin, M. (2023). Multimodal composition in second language education: A meta-analysis. *Language Teaching Research*, 27(6), 989–1012. https://doi.org/10.1177/13621688221124347

[17] Warschauer, M., & Grimes, D. (2011). Multimodal literacies in digital contexts. *Journal of Applied Linguistics*, 5(2), 121–140. https://doi.org/10.1558/japl.v5i2.121