Education Reform and Development, 2025, Volume 7, Issue 12
https://ojs.bbwpublisher.com/index.php/ERD
Online ISSN: 2652-5372
Print ISSN: 2652-5364

Research on the Innovation of Interactive Modes in Higher Vocational English Classrooms Driven by Multimodal Teaching Resources

Xinliang Cheng, Zhengyu Yan, Huiqin Bai

China Australia Business College of Shanxi, Jinzhong 030600, Shanxi, China

Copyright: © 2025 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

Abstract: To address problems such as insufficient interaction in higher vocational English classrooms and disconnection from vocational needs, this study focuses on the application value of multimodal teaching resources and explores their driving role in innovating classroom interactive modes. It first analyzes the importance of vocational scenario-based micro-courses, "text-image-practical operation" trimodal micro-course resources, and AI-enabled blended teaching closed loops in interactive innovation. Furthermore, it proposes targeted innovative strategies, including precise development of micro-courses, integration of trimodal resources, construction of AI closed loops, and optimization of evaluation systems. The research shows that multimodal teaching resources can effectively improve the quality of classroom interaction in higher vocational English, align with the talent training goals of vocational education, and provide practical references for the teaching reform of higher vocational English.

Keywords: Multimodal teaching resources; Higher vocational English; Classroom interactive modes; Vocational scenario-based micro-courses; AI empowerment

Online publication: December 31, 2025

1. Introduction

As a core public basic course in vocational education, higher vocational English undertakes the important mission of cultivating students' professional English application abilities. Currently, some higher vocational English classrooms face problems such as single interaction forms, low student participation, and disconnection between teaching content and vocational scenarios, which restrict the improvement of teaching quality. Relying on diversified presentation forms such as text, images, videos, and practical operations, multimodal teaching resources provide a new path for innovating classroom interactive modes [1]. Among them, the development of vocational scenario-based micro-courses, the construction of "text-image-practical operation" trimodal resources, and AI-enabled blended teaching closed loops have become key carriers connecting teaching with vocational needs. Based on this, this paper explores the importance and implementation strategies of multimodal teaching resources in the interactive innovation of higher vocational English classrooms, providing support for improving

teaching effectiveness.

2. The importance of multimodal teaching resources in innovating interactive modes of higher vocational English classrooms

2.1. Vocational scenario-based micro-courses lay a vocational-oriented foundation for interaction

The core goal of higher vocational English teaching is to cultivate students' English application abilities in vocational scenarios. However, traditional classroom interactions mostly revolve around textbook texts, which are disconnected from actual vocational situations, resulting in low student participation. Vocational scenario-based micro-courses, relying on real vocational scenarios, deeply integrate English knowledge with post skill requirements, laying a distinct vocational-oriented foundation for classroom interaction. By simulating typical post scenarios such as foreign trade negotiations, hotel services, and mechanical operation guidance, these micro-courses transform abstract English vocabulary and sentence patterns into concrete vocational behaviors, enabling students to clearly perceive the practical value of English learning [2]. In classroom interactions, students participate in micro-course-related role-playing and scenario simulation activities as "post practitioners," proactively using English to solve practical problems in vocational scenarios. This changes the previous state of passive knowledge acceptance and effectively improves the pertinence and effectiveness of interaction.

2.2. "Text-image-practical operation" trimodal resources enrich the multidimensional presentation of interaction

Traditional higher vocational English classroom interactions mostly rely on a single mode of text, with interaction forms limited to question-and-answer and reading, which are difficult to meet students' diverse cognitive needs. The "text-image-practical operation" trimodal micro-course resources integrate the accuracy of text information, the intuitiveness of image information, and the experience of practical operation links, providing a multidimensional presentation dimension for classroom interaction and making interaction forms more hierarchical and interesting. As the foundation, the text mode provides the core framework of English knowledge, such as key sentence patterns and professional vocabulary in vocational scenarios, ensuring the knowledge accuracy of interaction [3]; the image mode visualizes text content through pictures, animations, short videos, etc., reducing students' understanding difficulty. For example, when explaining mechanical operation English, the operation process is demonstrated through animations, combined with text explanations, enabling students to quickly master English expressions related to operation instructions; the practical operation mode guides students to transform learned knowledge into practical behaviors, such as completing the interpretation of English instructions for equipment operation and English expression of service processes according to the vocational scenario requirements in micro-courses [4]. During interaction, students can carry out diversified activities around trimodal resources, such as organizing information by combining text and images, verifying learning results through practical operations, and sharing practical experience through group discussions, forming an interactive closed loop of "cognition-practice-reflection".

2.3. AI-enabled blended teaching closed loops improve the precision of interaction

In traditional classroom interactions, it is difficult for teachers to fully grasp each student's learning status and interaction effects, resulting in a lack of precision in interaction guidance. The integration of multimodal teaching

resources and AI technology constructs a blended teaching closed loop of "pre-class preview - in-class interaction - post-class consolidation", providing technical support for improving the precision of interaction. Before class, the AI platform pushes multimodal preview resources, such as vocational scenario-based micro-courses and trimodal exercises, and records students' preview progress, answer accuracy, and other information through data monitoring to accurately identify students' knowledge weaknesses [5]. During in-class interaction, based on the preview data fed back by AI, teachers design targeted interactive tasks. For example, for the knowledge point of "cross-border e-commerce logistics English expression" that students generally fail to master, group cooperation activities such as interpreting logistics documents and simulating logistics communication scenarios are carried out. At the same time, AI captures data such as students' language expression and task completion during interaction in real time, analyzes indicators such as students' pronunciation accuracy and participation through speech recognition and image recognition technologies, and timely feeds back interaction effects to teachers [6]. After class, AI pushes personalized consolidation resources according to in-class interaction data, such as special pronunciation correction micro-courses for pronunciation problems and extended scenario exercises for insufficient scenario expression, and provides Q&A guidance for students through online interaction communities, forming a continuous optimization closed loop of interaction effects.

3. Innovation strategies of interactive modes in higher vocational English classrooms driven by multimodal teaching resources

3.1. Precisely develop vocational scenario-based micro-courses to build an interactive content support system

Vocational scenario-based micro-courses are the core carriers of interaction driven by multimodal resources, and their development quality directly determines the interaction effect. The principle of "post demand orientation and precise content adaptation" should be adhered to. First, conduct post research, form a development team composed of front-line enterprise experts and English teachers, and sort out core vocational scenarios and English ability requirements corresponding to different majors through interviews and on-site investigations. For example, focus on scenarios such as inquiry, quotation, and contract signing for international trade majors, and form a "major-scenario-English knowledge point" correspondence list [7]. Second, optimize micro-course design. Each micro-course focuses on one core vocational scenario, with a duration of 5-8 minutes, and integrates "text-image-practical operation" trimodal elements: the text part refines core vocabulary, sentence patterns, and pragmatic rules in the scenario, with concise and standardized language expression; the image part presents visual information such as scenario environments and operation processes through on-site shooting and animation production [8]; the practical operation part designs 1-2 key operation links, such as simulating the filling of customs declaration forms and demonstrating English communication scripts for international trade contract signing. At the same time, set up interaction guidance links in micro-courses, such as questions like "How to respond to customers' price doubts in this scenario" and "Try to describe the next operation steps in English", reserving interfaces for classroom interaction. Finally, establish a micro-course resource library, store resources by major characteristics and teaching progress, and update them regularly according to changes in post demands to ensure the timeliness and applicability of resources, providing continuous and stable content support for classroom interaction [9].

3.2. Integrate "Text-image-practical operation" trimodal resources to design hierarchical interactive activities

Based on the characteristics of "text-image-practical operation" trimodal resources, design hierarchical and progressive classroom interactive activities to guide students to deepen from knowledge cognition to practical application, and improve the participation and depth of interaction. Basic cognitive interaction focuses on the integration of text and images, with the goal of knowledge understanding. Teachers can use multimedia equipment to display text and image content in trimodal resources and design interactive tasks such as "information matching" and "key point extraction". For example, when explaining the "product introduction" scenario, present product manual texts and product structure diagrams, and let students work in groups to match professional vocabulary in the text with corresponding parts in the image and extract the core selling points of the product in English [10]. Through this level of interaction, students can consolidate their language foundation and understand the semantic connection between modes. Ability application interaction focuses on the integration of practical operation modes, with the goal of skill application. Combined with the practical operation links in micro-courses, carry out interactive activities such as "scenario simulation" and "role-playing". For example, in hotel English courses, provide text scripts, service process images, and practical operation requirements for guest room service scenarios. Students are divided into groups to play waiters and guests, completing practical interactions such as check-in guidance and demand response. Teachers record the interaction process through on-site observation and video recording, and organize review and comments after class. Thinking expansion interaction focuses on the comprehensive application of multimodal resources, with the goal of cultivating innovative abilities [11]. Design open interactive tasks such as "problem solving" and "scheme design". For example, provide text cases of foreign trade order delays and logistics tracking images, and let students work in groups to develop English communication schemes, solving problems through simulated phone calls and writing apology letters, cultivating students' comprehensive application and innovative thinking abilities. Hierarchical interactive activities take into account the needs of students at different levels, enabling each student to improve in interaction.

3.3. Build an AI-enabled blended teaching closed loop to realize dynamic optimization of interaction

Relying on AI technology, build an integrated blended teaching closed loop of "pre-class - in-class - post-class", and realize precise design, real-time feedback, and dynamic optimization of interaction through data-driven methods. In the pre-class preview stage, the AI platform pushes adaptive multimodal preview resources according to teaching goals, including vocational scenario-based micro-courses, trimodal exercises, and online preview questionnaires. After students complete the preview, AI automatically analyzes the preview data, generates knowledge mastery reports for the whole class and individual students, and identifies common weaknesses of the class and personalized needs of students. Based on the reports, teachers design targeted in-class interaction plans. For example, for the knowledge point of "past participles and present participles as attributes" that the class generally confuses, design multimodal comparative analysis interactions; for individual students with unsmooth expression in vocational scenarios, arrange one-on-one scenario simulation interactions [12]. In the in-class interaction stage, carry out diversified interactive activities with the help of AI interactive teaching platforms, such as English pronunciation competitions through speech recognition technology, evaluation of the standardization of practical operations through image recognition technology, and collection of students' interactive answers through real-time answering systems. AI captures interaction data in real time, such as participation, answer accuracy, and pronunciation accuracy, and generates an in-class interaction data dashboard

immediately. Teachers can dynamically adjust the interaction rhythm and difficulty according to the data. For example, when finding a group has problems with non-standard expression in scenario simulation interaction, pause immediately and provide targeted guidance through replaying micro-course clips [13]. In the post-class consolidation stage, AI pushes personalized multimodal consolidation resources to students according to pre-class preview and in-class interaction data, such as special pronunciation correction micro-courses and speech practice tasks for students with weak pronunciation, and extended scenario simulation materials for students with insufficient scenario application.

3.4. Improve the multidimensional evaluation system to ensure the implementation effect of interactive modes

Construct a multidimensional evaluation system adapted to the interactive mode driven by multimodal resources, and ensure the implementation effect of teaching reform through comprehensive and objective evaluation. The evaluation subjects adopt a diversified model of "teacher evaluation + student self-evaluation + group mutual evaluation + AI evaluation" to ensure comprehensiveness. Teacher evaluation focuses on knowledge application, skill performance, and professional literacy during interaction, and conducts comprehensive scoring based on classroom observation and interaction records; student self-evaluation focuses on self-reflection on participation, learning gains, and shortcomings during interaction, completed by filling out self-evaluation questionnaires [14]; group mutual evaluation centers on team collaboration and task contribution, with scores jointly given by group members; AI evaluation relies on technical advantages to conduct objective scoring on data-based indicators such as students' pronunciation accuracy, interactive task completion efficiency, and online interaction activity. The evaluation content covers three core dimensions: knowledge mastery, ability improvement, and literacy development. The knowledge mastery dimension focuses on students' mastery of English vocabulary, sentence patterns, and pragmatic rules, evaluated through in-class answers and post-class tests; the ability improvement dimension focuses on professional English application abilities, evaluated based on performance in scenario simulation and practical interaction; the literacy development dimension focuses on professional literacy such as team collaboration, communication, and professional responsibility, evaluated through behavioral performance during interaction. The evaluation results adopt a combination of "process evaluation + summative evaluation", with process evaluation accounting for 60%, covering pre-class preview, in-class interaction, and post-class consolidation; summative evaluation accounts for 40%, mainly based on comprehensive application assessment of English in vocational scenarios [15]. At the same time, establish an evaluation feedback mechanism, regularly feed back evaluation results to students, put forward improvement suggestions for shortcomings, and guide students to actively adjust their interaction participation methods, forming a continuous improvement mechanism of "evaluation - feedback - improvement".

4. Conclusion

Multimodal teaching resources provide important support for innovating the interactive modes of higher vocational English classrooms, and their importance in vocational orientation guidance, interaction dimension enrichment, precision improvement, and collaborative education is increasingly prominent. Through strategies such as precisely developing vocational scenario-based micro-courses, designing trimodal hierarchical interactive activities, building AI-enabled blended teaching closed loops, and improving the multidimensional evaluation system, the problems of traditional classroom interaction can be effectively solved, and students' classroom

participation and professional English application abilities can be improved. The interactive innovation mode constructed in this study conforms to the type characteristics of higher vocational education, provides a practical path for the teaching reform of higher vocational English, and promotes the further improvement of higher vocational English teaching quality.

Funding

2025 Research Project on the High-Quality Development of Vocational Foreign Language Education by the Advisory Committee of Foreign Language Teaching in Vocational Education, Ministry of Education, P.R. China, "Research on the Development of Multimodal Digital Resources and Blended Teaching Practice for Vocational Scenarios in English Teaching of Private Higher Vocational Colleges" (Project No.: WYJZW-202513342SX); 2024 Special Project on Foreign Language Education Reform in Vocational Colleges by Foreign Language Education Committee of the Chinese Society for Technical and Vocational Education (CSTVE), "Research on the Optimization Strategies for the Professional Development of 'Dual-Qualified' Foreign Language Teachers in Private Higher Vocational Colleges from the Perspective of New Quality Productive Forces" (Project No.: WYW2024C16)

Disclosure statement

The authors declare no conflict of interest.

References

[1] Niu L, 2024, Research on Innovative Practice of Multimodal Teaching of Higher Vocational English Under the Internet Environment. Overseas English, 19: 215–217.

[2] Hu J, Wang Y, 2024, Multimodal Teaching Strategies for Ideological and Political Education in College English Courses Under the Background of Digital Transformation. Teaching of Forestry Region, 8: 29–32.

[3] Tian Q, 2024, Research on Multimodal Teaching of Higher Vocational English Under the Background of Big Data Technology. Journal of Lanzhou Vocational Technical College, 40(4): 37–39.

[4] Xu M, 2023, Research on the Path to Improve Higher Vocational English Teachers' Multimodal Teaching Competence. Education Informatization Forum, 8: 45–47.

[5] Zhang W, 2023, Exploration of Multimodal Theory-Based Teaching Strategies for Higher Vocational English. The Road to Success, 19: 113–116.

[6] Hu T, Liao M, 2023, Research on Multimodal Teaching of Higher Vocational English from the Perspective of the Internet. Guangdong Vocational and Technical Education and Research, 2: 71–73.

[7] Fang F, 2022, Research on Multimodal Teaching Mode of Higher Vocational English Under the Background of Curriculum Ideological and Political Education. Modern Vocational Education, 40: 128–130.

[8] Shen R, 2020, Practical Exploration of Multimodal Teaching Mode in Higher Vocational English Teaching. China Journal of Multimedia and Network Teaching (Mid-Monthly), 12: 42–43 + 46.

[9] Luo W, Tang N, Liu X, 2020, Construction of Multimodal Teaching Mode for Higher Vocational English Speaking Courses. Hebei Vocational Education, 4(1): 64–66.

[10] Huang D, Zeng M, 2019, Multimodal Teaching Mode of Higher Vocational Tourism English Based on Online

Teaching Platform—Taking “Hotel Situational English” as an Example. Journal of Xiamen City University, 21(4): 62–67.

[11] Hu E, 2019, Research on Multimodal Teaching of Higher Vocational English in the Era of Big Data. Neijiang Science & Technology, 40(7): 131–132.

[12] Cheng J, 2019, Research on Multimodal Teaching Mode of Higher Vocational English Based on Flipped Classroom. China Journal of Multimedia and Network Teaching (Mid-Monthly), 3: 31–32.

[13] Zhang L, 2019, Research on the Construction of Teaching Resource Platforms Under the Network Environment—Taking the Construction of English Curriculum Resource Library as an Example. Think Tank Era, 2: 126 + 128.

[14] Zhou X, 2018, Research on the Application of Flipped Classroom Teaching Mode in Multimodal Teaching of Higher Vocational English. Liberal Arts Enthusiast (Education and Teaching), 12: 104–105.

[15] Liao S, 2014, Multimodal Interactive Classroom Teaching of College English in Higher Vocational Colleges Under the Multimedia Environment. Education and Vocation, 33: 148–149.

Publisher’s note

Bio-Byword Scientific Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.