

Research on the Theory and Practice of AI-enabled Digitalization in College English Listening and Speaking Courses

Yafei Pang

Foshan University, Foshan, 528000, China

Email: Pangyafei@fosu.edu.cn

How to cite this paper: Pang, Y. F. (2025). Research on the Theory and Practice of AI-enabled Digitalization in College English Listening and Speaking Courses. *International Journal of Educational Development*, 2(1), 39-49. ISSN Print: 3104-4263; ISSN Online: 3104-4271. <https://doi.org/10.63313/IJED.9033>
Published: 2025-12-22

Copyright © 2025 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/>



Abstract

This study focuses on the digital-intelligent transformation of college English listening and speaking courses, aiming to address core pain points of traditional teaching such as limited opportunities to speak, delayed feedback, outdated materials, and insufficient scenarios. Adopting a control experiment method, two parallel classes of first-year non-English majors from our university were selected as research objects: the experimental class (52 students) utilized the *Chaoxing* Platform (a typical carrier of AI-enabled digital education), integrating innovative teaching models including AI-generated CET-4 listening materials, human-computer interactive oral training, and language practices such as PPT presentations and video production; the control class (50 students) adopted traditional listening and speaking teaching methods. The research cycle lasted one semester, with data collected through final exam paper analysis and face-to-face interviews to systematically explore the impact of AI-driven teaching on students' listening and speaking abilities and comprehensive language proficiency. Results indicate that students in the experimental class outperformed their peers in the control class in terms of listening interest and oral expression authenticity. Meanwhile, our findings indicate that AI-enabled Digitalization teaching boosts not only English listening and speaking proficiency, but also other key skills including reading, translation and writing. However, problems such as students' excessive reliance on AI and weakened willingness to think independently were also identified. This study verifies the effectiveness of the teaching model integrating AI technology with the *Chaoxing* Platform in cultivating listening and speaking skills, providing a replicable practical paradigm for the intelligent digitalization transformation of college English teaching and targeted suggestions for mitigating technology application risks.

Keywords

College English; Listening and Speaking Courses; AI-enabled Digitalization; Theory and Practice; Human-Computer Interaction

1. Research Background

1.1. Practical Dilemmas in College English Listening and Speaking Teaching

As a core public course in higher education, college English undertakes the important mission of cultivating students' cross-cultural communication competence and language application abilities. Traditional listening and speaking teaching has long faced multiple bottlenecks: Firstly, the "difficulty in speaking up" is prominent in oral teaching. The large-class teaching model makes it difficult for teachers to provide personalized guidance, and students avoid expression due to anxiety. Moreover, there is a lack of real language communication scenarios-limited role-plays are mostly confined to daily topics, with a serious shortage of high-level scenario training such as academic discussions and cross-cultural negotiations. Secondly, listening teaching suffers from outdated materials and poor difficulty adaptability. The content of news listening in the College English Test Band 4 (CET-4) is updated slowly, failing to meet students' needs for exam preparation and practical language application. Thirdly, teaching feedback is delayed. Pronunciation errors and inappropriate expressions in oral practice are difficult to correct in real time, and error analysis in listening training lacks targeted interpretation, leading to low learning efficiency. Fourthly, the teaching model is rigid and simplistic, over-reliant on textbooks and classroom lectures, lacking immersive and interactive learning experiences, which makes it difficult to stimulate students' learning initiative.

1.2. Era Opportunities of AI-enabled Digital Education

AI-enabled digital Education, centered on big data and artificial intelligence technologies, realizes the digital reconstruction and intelligent upgrading of the entire teaching chain through "data-driven decision-making, intelligent process optimization, and resource collaborative sharing," covering four dimensions: resource digitization, teaching intelligence, evaluation precision, and management efficiency (Zhu Huidong, 2025)[1]. We argue that the construction of AI-enabled digitalization in college English listening and speaking courses can effectively address the current problems. Artificial intelligence technologies, with core functions such as big data, natural language processing (NLP), automatic speech recognition (ASR), and text-to-speech (TTS), provide technical support for the innovation of college English listening and speaking teaching (Zhang Yanlu, 2019[2]; Pang, Y., He, Y., & Marlow, D., 2025[3]; Pang, Y., Marlow, D., & He, Y., 2025[4]). Large language models (LLMs) such as ChatGPT and Doubao possess multimodal interaction, context understanding, and real-time generation capabilities, enabling them to simulate real dialogue scenarios, accurately identify language errors, and provide immediate feedback. Smart teaching platforms like Chaoxing have realized functions such as teaching resource integration, learning situation data analysis, and personalized resource recommendation, providing a stable carrier for human-machine collaborative teaching. Domestic

and foreign studies have confirmed that AI technology can effectively reduce students' oral anxiety (Dizon & Tang, 2020[5]; Chen et al., 2022[6]), increase speaking frequency and language accuracy. Its application value in listening material adaptation (Pang, Y., & Marlow, D., 2025[7]), oral interactive training (Kong Lei, 2024)[8], and intelligent assessment (Jeon & Lee, 2023)[9] has been initially verified.

2. Theoretical Exploration

2.1. Second Language Acquisition Theories

2.1.1. Input Hypothesis and "i+1" Principle

As a core theory of second language acquisition, the Input Hypothesis emphasizes that comprehensible input is the key to language acquisition (Krashen, 1985)[10]. Its "i+1" principle states that language input should be slightly higher than the learner's current level (i), ensuring that learners can understand the materials based on existing knowledge while promoting language ability improvement through moderate challenges. AI technology perfectly aligns with this principle: large language models can adapt news texts, cultural materials, etc., into listening materials suitable for CET-4 level, precisely controlling vocabulary difficulty and sentence complexity to achieve personalized "i+1" input. This meets the learning needs of students with different foundations and addresses the "one-size-fits-all" problem in traditional teaching.

2.1.2. Output Hypothesis

The Output Hypothesis argues that language acquisition requires not only sufficient input but also meaningful language output (Swain, 1995)[11]. Oral output activities encourage learners to actively organize language knowledge, identify their own deficiencies, and thereby achieve language internalization. AI technology provides efficient support for output training: intelligent chatbots and virtual language partners can offer all-weather, low-pressure oral interaction scenarios with real-time feedback on expression errors; human-computer interaction helps students optimize language output and strengthen knowledge internalization.

2.2. TPACK Theory

Technological Pedagogical Content Knowledge (TPACK) theory is the core methodology for AI-empowered teaching. Its essence lies in the organic integration of teaching content, teaching methods, and technical tools to form a dynamic teaching knowledge system (Wang, Y., et al., 2024[12]; Zhu, Y., 2025[13]). In this study, the teaching content focuses on core listening and speaking skills; the teaching method adopts a full-process interactive model of "pre-class - in-class - post-class"; the technical tools, centered on the Chaoxing Platform intelligent agent, integrate functions such as AI listening material generation, human-computer interaction, and intelligent revision. TPACK theory emphasizes that technology serves teaching objectives

and avoids technology abuse. This concept guides the construction of a precise adaptation relationship of "AI technology - listening and speaking teaching - ability improvement" in this study, ensuring that technology application meets actual teaching needs.

2.3. Situated Learning Theory

Situated Learning Theory holds that learning is a process in which learners interact with the environment and construct knowledge in real contexts. The cultivation of English listening and speaking abilities is inseparable from real language use scenarios, which are difficult to replicate in traditional teaching. AI technology can construct immersive language communication scenarios: for example, simulating cross-cultural negotiations, academic seminars, daily dialogues, etc., allowing students to practice oral expression and listening comprehension abilities through interaction. The group collaboration function of the Chaoxing Platform can be combined with AI-generated thematic tasks to realize online-offline integrated situational practice, helping students transform language knowledge into practical application abilities.

3. Application of Theory to Practice

Based on the theoretical exploration and relevant literature review of this research, the study applies theories to practice.

3.1. Experimental Design

3.1.1. Research Objects

Two parallel classes of first-year non-English majors (2024 intake) from the university were selected as research objects: 52 students in the experimental class and 50 in the control class. The consistency of the two groups' basic English proficiency was verified through the entrance English proficiency test (including four parts: speaking, listening, reading, and writing, with a full score of 100). The average score of the experimental class was 68.3, and that of the control class was 67.9. An independent samples t-test showed $t=0.523$, $p>0.05$, indicating no significant statistical difference and ensuring experimental comparability.

3.1.2. Experimental Cycle

One semester (16 weeks), with 3 class hours of college English listening and speaking courses and 3 class hours of independent learning for comprehensive English courses per week.

3.1.3. Experimental Variables

Independent variable: Teaching model. The experimental class adopted a mixed teaching model of "Chaoxing Platform+AI tools"; the control class adopted the tradi-

tional listening and speaking teaching model (textbook teaching, in-class listening practice, group discussion, and teacher revision of assignments).

Dependent variables: Students' English listening and speaking abilities (with final exam listening and speaking scores as core indicators), learning interest, and comprehensive language proficiency (reading, translation, and writing scores).

Controlled variables: Textbook (New College English: An Integrated Course 1 published by Foreign Language Teaching and Research Press, 2020), teaching teachers, class hour arrangement, and final exam paper difficulty.

3.2. Teaching Implementation Path of the Experimental Class

Relying on the *Chaoxing* Platform intelligent agent, a full-process teaching model of "AI-empowered input - human-computer interactive output - practical application enhancement - intelligent evaluation feedback" was constructed, specifically divided into three modules: listening teaching, oral teaching, and comprehensive practice.

3.2.1. Listening Teaching: AI-Generated CET-4-Oriented Personalized Material System

With CET-4 news listening as the core objective, real-time news materials were adapted using large language models based on Krashen's "i+1" principle.

Pre-class: Teachers designed prompts through large language models (e.g., *Doubao*) to adapt recent news from authoritative media such as China Daily (e.g., international events, technological developments) into CET-4-level listening materials, slightly more difficult than real exam questions (adding professional terms, complex proper nouns at the vocabulary level, and multi-dimensional modified long sentences with implicit logical connections at the sentence structure level).

In-class: Firstly, the audio playback function of the *Chaoxing* Platform was used for listening training, supporting speed adjustment and repeated playback; secondly, AI analysis tools (e.g., *DeepSeek*) were used for syntactic analysis and cultural expansion of long and difficult sentences in the listening materials to help students understand language structures and contextual meanings; finally, group discussions were organized to conduct critical thinking around the themes of the listening materials (e.g., cross-cultural communication, technological ethics) to deepen listening comprehension.

Post-class: The *Chaoxing* Platform pushed targeted consolidation exercises, including error re-practice and extended training on similar topics; AI tools (e.g., *Hailuo AI*) allowed students to upload their own listening recordings for speech recognition and accuracy analysis, providing feedback on the impact of pronunciation problems on listening comprehension (e.g., information loss caused by liaison and weak pronunciation).

3.2.2. Oral Teaching: Situational Interaction and Practice with Human-Computer Collaboration

Centered on "intelligent language partners + real tasks" and combined with the interactive functions of the *Chaoxing* Platform, students' oral expression fluency, authenticity, and application abilities were improved.

Pre-class: The *Chaoxing* Platform pushed oral task themes (e.g., "Campus Life," "Cross-Cultural Communication," "Academic Discussion"). AI intelligent language partners (based on the built-in model of the *Chaoxing* Platform) provided personalized preview guidance: for example, for the theme of "Academic Discussion," relevant academic vocabulary collocations and sentence pattern examples were generated; students could conduct preview dialogues with AI language partners, which provided real-time correction of pronunciation errors and grammatical problems, reducing in-class expression anxiety.

In-class: A "human-computer interaction + interpersonal collaboration" model was adopted. Firstly, students were divided into groups to conduct situational simulation dialogues with AI language partners (e.g., simulating multinational company interviews, academic seminar debates). AI played different roles (e.g., interviewer, challenger) to promote in-depth expression through questioning, refutation, etc.; secondly, relying on the group collaboration function of the *Chaoxing* Platform, students presented dialogue results in PPT format, and AI provided real-time feedback on language expression issues in the PPT (e.g., single sentence structure, inappropriate vocabulary); finally, teachers guided class-wide comments, focusing on the authenticity and logical coherence of oral expression.

Post-class: Video production tasks were assigned, requiring students to create bilingual short videos around unit themes (e.g., "Good Interpersonal Relationships"). AI tools supported script optimization (supplementing cultural details and authentic expressions) and speech synthesis (adjusting speech rate and accent). After students uploaded the videos to the *Chaoxing* Platform, AI scored them from three dimensions: language accuracy, expression fluency, and content completeness, while opening channels for peer evaluation and teacher comments.

3.3. Teaching Implementation Path of the Control Class

The traditional listening and speaking teaching model was adopted to form a contrast with the experimental class:

Listening teaching: Textbook supporting listening materials and previous CET-4 real exam listening materials were used. After playing the audio in class, answers and difficult points were explained. Paper-based listening exercises were assigned after class, which were uniformly revised by teachers before feedback.

Oral teaching: Focused on textbook dialogue practice and group discussion. Teachers explained pronunciation and grammar, organized role-play activities (without AI language partner preview), and students presented results orally. Teachers provided overall comments without personalized feedback.

Post-class tasks: Written assignments (e.g., dialogue script writing, compositions) were assigned, which were manually revised by teachers with a feedback cycle of

approximately 3-5 days; there were no video production or AI-assisted optimization links.

4. Discussion and Analysis

4.1. Final Exam Paper Analysis

The university's college English final exam paper consists of six parts: Oral Test, Listening Comprehension, Vocabulary and Structure, Reading Comprehension, C-E Translation, and Writing, with a total score of 100.

4.1.1. Comparison of Listening and Speaking Ability Scores

The listening and speaking part of the final exam is worth 30 points, including listening comprehension (15 points) and oral expression (15 points). The score statistics of the two groups are as follows (Table 1):

Table 1. Comparison of Listening and Speaking Scores between the Experimental Class and the Control Class.

Class	Number of Students	Average Listening Comprehension Score (Full Score: 15)	Average Total Oral and Listening Score (Full Score: 30)	Average Oral Score (Full Score: 15)	Pass Rate (≥ 21 points)
Experimental Class	52	11.8	24.1	12.3	88.5%
Control Class	50	10.6	21.9	11.3	62.0%

Independent samples t-test results showed significant statistical differences in listening and speaking abilities between the experimental class and the control class ($p < 0.001$): the average listening score of the experimental class reached 11.8 points, 1.2 points higher than that of the control class (10.6 points) ($t = 6.32$); the average oral score was 12.3 points, 1.0 point higher than that of the control class (11.3 points) ($t = 5.78$); the average total listening and speaking score was 24.1 points, 2.2 points higher than that of the control class (21.9 points) ($t = 7.15$); the pass rate (≥ 21 points) reached 88.5%, 26.5 percentage points higher than that of the control class (62.0%).

Specifically, the experimental class students achieved a significant improvement in the accuracy of detail questions and inference questions in listening, which is closely related to the special training using AI-generated high-simulation CET-4 listening materials; in terms of oral expression, the usage rate of authentic expressions (e.g., prepositional collocations, fixed sentence patterns) of the experimental class students was 32% higher than that of the control class, and the pronunciation error rate was reduced by 28%. However, in daily impromptu oral scenarios (e.g., random topic Q&A), the average score of the experimental class was only 1.2 points higher than that of the control class, with a relatively limited improvement range. This indicates that the cultivation of impromptu oral ability requires longer-term systematic training and practical accumulation.

4.1.2. Comparison of Comprehensive Language Proficiency Scores

The Vocabulary and Structure, Reading Comprehension, C-E Translation, and Writ-

ing of the final exam are worth 70 points. The score statistics of the two groups are as follows (Table 2):

Table 2. Comparison of Other Final Exam Scores between the Experimental Class and the Control Class

Class	Number of Students	Average Vocabulary and Structure(Full Score: 15)	Average Reading Score (Full Score: 30)	Average Translation Score (Full Score: 10)	Average Writing Score (Full Score: 15)	Average Comprehensive Score (Full Score: 70)
Experimental Class	52	12	21	8	12.5	53.5
Control Class	50	11	19	7	10	47

The average comprehensive score of the experimental class (53.5 points) was 6.5 points higher than that of the control class (47 points) ($t=5.93$, $p<0.001$), indicating that AI-driven listening and speaking teaching has a positive transfer effect on reading, translation, and writing abilities. This conclusion is fully supported by the differences in sub-item scores between the two classes: the experimental class outperformed the control class in vocabulary and structure (12 points vs. 11 points), reading (21 points vs. 19 points), translation (8 points vs. 7 points), and writing (12.5 points vs. 10 points). This is mainly due to: AI-pushed listening materials and vocabulary collocation training strengthened students' vocabulary accumulation and grammar mastery, which directly contributed to the improvement of their vocabulary and structure scores; AI revision in oral output and script writing helped students optimize language expression, thereby enhancing the accuracy and authenticity of their reading comprehension, writing, and translation, as reflected in the higher scores of these three sub-items in the experimental class.

4.2. Analysis of Face-to-Face Interview Results

Through face-to-face interviews with 20 students from the experimental class and 15 students from the control class, combined with teachers' teaching observations, the advantages and existing problems of AI-enabled digital teaching were summarized:

4.2.1. Significant Advantages

Enhanced interest in listening learning: 85% of the experimental class students believed that AI-generated real-time news listening materials were more attractive than traditional textbook materials, "closer to real life" and "moderately difficult with challenges," which could stimulate active learning willingness; in contrast, 60% of the control class students reported that "listening materials are outdated" and "boring," leading to low learning motivation.

Increased confidence and authenticity in oral expression: 78% of the experimental class students stated that the low-pressure interaction with AI language partners reduced oral anxiety, "daring to speak without fear of making mistakes"; at the same time, the authentic expression suggestions provided by AI made their expressions "more like native speakers." In contrast, control class students generally believed

that "oral practice lacks feedback" and "do not know if they are speaking correctly." Improved learning efficiency: Experimental class students reported that the personalized resource recommendation and AI real-time feedback on the Chaoxing Platform saved a lot of time in searching for materials and waiting for revision, "able to quickly identify their own problems and make targeted improvements"; teachers stated that AI-assisted revision and learning situation analysis reduced their workload, allowing them to focus more on cultivating high-level abilities.

4.2.2. Existing Problems

Excessive reliance on AI and weakened independent thinking ability: 70% of the experimental class students admitted that when writing oral scripts and optimizing PPT texts, they would directly copy AI's revision suggestions, "being too lazy to think about expressions independently"; some students would immediately use AI to look up new words when encountering them in listening practice, lacking training in guessing word meanings from context, leading to a decline in independent problem-solving ability.

Limited improvement in impromptu oral ability: Interviews found that experimental class students performed well in structured oral tasks (e.g., PPT presentations, video recording) but still faced problems such as unclear thinking and unsmooth expression in impromptu expression of random topics. Students reflected that "being accustomed to AI's guidance and feedback, they would feel nervous when facing unprepared topics," indicating that AI-assisted structured training has not been fully transformed into impromptu expression ability.

Technology adaptability issues: Some students reported that the accuracy of AI in recognizing non-standard accents was insufficient, "sometimes their pronunciation is correct, but AI still judges it as wrong"; the AI language partner of the Chaoxing Platform has limited understanding ability in complex contexts, "unable to respond to discussions on some in-depth topics."

4.3. Problem Attribution and Improvement Suggestions

4.3.1. Problem Attribution

Teaching design level: Excessive emphasis on AI's auxiliary functions, lack of targeted training on students' independent thinking ability, such as insufficient design of open tasks, leading students to form a thinking pattern of "relying on AI to obtain standard answers."

Student cognition level: Some students regard AI as an "all-purpose tool," ignoring that the essence of language learning is "active construction and practice," and lacking in-depth understanding and internalization of language knowledge.

Technology itself level: Current AI models still have limitations in in-depth context understanding and non-standard accent recognition, making it difficult to fully replace teachers' personalized guidance and emotional communication.

4.3.2. Improvement Suggestions

Balance AI application and independent thinking in AI-enabled digital teaching: Reduce AI's function of directly providing answers, increase open and inquiry-based tasks, such as designing oral tasks of "AI raising questions, students independently searching for information and demonstrating"; require students to mark revision reasons after using AI to revise texts, strengthening the understanding of language knowledge.

Strengthen teacher guidance and give play to the advantages of human-machine collaboration: Based on the learning situation data of the Chaoxing Platform, teachers should accurately identify students' dependence points, such as designing word meaning guessing training for students who overuse AI to look up words; in oral teaching, increase impromptu topic training to guide students to express independently without AI assistance.

5. Conclusion

Through a quasi-experiment, this study verifies the significant effectiveness of the AI-enabled digital teaching model for college English listening and speaking courses in cultivating students' English listening and speaking abilities. Students in the experimental class outperformed those in the control class in terms of listening and speaking scores, learning interest, and comprehensive language proficiency. AI-enabled digital teaching effectively addresses the problems of outdated materials, delayed feedback, and insufficient scenarios in traditional listening and speaking teaching: AI-generated personalized listening materials improve learning pertinence, human-computer interactive oral training reduces expression anxiety, and intelligent evaluation on the Chaoxing Platform realizes real-time feedback and precise teaching. At the same time, the study also identifies problems such as excessive reliance on AI and limited improvement in impromptu oral ability, which need to be solved by optimizing teaching design, strengthening teacher guidance, and improving technology adaptability.

Future research can further extend the experimental cycle to explore the impact of AI-enabled digital teaching on students' long-term language ability development; at the same time, VR/AR technology can be integrated to enrich immersive teaching scenarios, and regional characteristic resources such as Lingnan culture can be incorporated to realize the coordinated development of language skill training and cultural confidence cultivation. In conclusion, AI-enabled digital classrooms have brought a paradigm innovation to college English listening and speaking teaching. However, only by adhering to the principles of "technology serving teaching, teachers leading teaching, and students learning independently" can its value be fully exerted, cultivating students' comprehensive English application abilities that meet the needs of the times.

Acknowledgement

This paper is a phased achievement of the 2025 Guangdong Provincial Education Science Planning Project (Scientific Research Topics on Higher Education): Research on the Construction of AI-enabled Digital Foreign Language Curricula for the Integration of Excellent Local Traditional Culture and Its International Dissemination—Taking Lingnan Culture as an Example (Project No.: 2025GXJK0538).

References

- [1] Zhu, H.D. (2025) Strategies and Practice of Integrating Digital-Intelligent Empowerment into Foreign Language Courses and Regional Culture in Application-Oriented Undergraduate Universities. *Modern English*, 9, 76-78.
- [2] Zhang, Y.L. (2019) Research on Artificial Intelligence Promoting the Reform of College English Teaching. *Education Modernization*, 6, 130-131.
- [3] Pang, Y.F., He, Y. and Marlow, D. (2025) Employing Large Language Models in College English Teaching: Insights from China. *International Journal of Social Sciences, Language and Linguistics*, 5, 16-21.
- [4] Pang, Y.F., Marlow, D. and He, Y. (2025) Research on Empowering College English Teaching with Artificial Intelligence. *Education and Social Work*, 1, 131-139.
- [5] Dizon, G. and Tang, D. (2020) Intelligent personal assistants for autonomous second language learning. *JALT CALL Journal*, 16, 107-120.
- [6] Chen, C., Koong, C. and Liao, C. (2022) Influences of integrating dynamic assessment into a speech recognition learning design. *Educational Technology & Society*, 25, 1-14.
- [7] Pang, Y.F. and Marlow, D. (2025) Application of Artificial Intelligence in College English Listening Teaching. *International Journal of Social Sciences, Language and Linguistics*, 10, 18-24.
- [8] Kong, L. (2024) Application of Generative Artificial Intelligence in Foreign Language Major Teaching: A Case Study of College Critical English: Intensive Reading. *Foreign Language Education Research Frontiers*, 7, 11-18+90.
- [9] Jeon, J. and Lee, S. (2023) Large language models in education: A focus on the complementary relationship between human teachers and ChatGPT. *Education and Information Technologies*, 28, 58-60.
- [10] Krashen, S.D. (1985) *The Input Hypothesis: Issues and Implications*. London: Longman.
- [11] Swain, M. (1995) Three Functions of Output in Second Language Learning. In: Cook, G. and Seidlhofer, B. (Eds.) *Principle and Practice in Applied Linguistics: Studies in Honor of Widdowson, H. G.* Oxford: Oxford University Press.
- [12] Wang, Y., Du, Y., Zhao, P. et al. (2024) Research on College English Ideological and Political Education Model Based on TPACK Theory. *Overseas English*, 20, 149-151.
- [13] Zhu, Y. (2025) Construction and Practice of TPACK-Based Smart Classroom Teaching Model for College English. *Journal of Wuxi Vocational and Technical College*, 4, 35-41.