ISRG Journal of Education, Humanities and Literature (ISRGJEHL)

ISRG PUBLISHERS

Abbreviated Key Title: ISRG J Edu Humanit Lit

ISSN: 2584-2544 (Online)

Journal homepage: https://isrgpublishers.com/isrgjehl/

Volume – III Issue – VI (November-December) 2026

Frequency: Bimonthly



AI-Enabled Instruction for L2 Speaking: Effects on Learners' Oral Proficiency

Yafei Pang[1], David Marlow[2]

[1] Foshan University, Foshan, Guangdong, China.

[2] University of South Carolina Upstate, Spartanburg, South Carolina, USA

| **Received:** 05.11.2026 | **Accepted:** 07.11.2026 | **Published:** 11.11.2026

***Corresponding author:** Yafei Pang

Abstract

This study focuses on the application of artificial intelligence (AI) technology in second language (L2) oral instruction, aiming to address the long-standing practical difficulties in L2 speaking pedagogy, including large class sizes which result in limited opportunities for speaking and delayed feedback from instructors, lack of authentic communicative contexts, and high levels of speaking anxiety. This study adopts a quasi-experimental design. Two parallel classes of non-English majors from related disciplines in the first year at the author's university were selected as research subjects. The experimental class (28 students, 2025 intake, Mathematics and Applied Mathematics) received an AI-driven innovative teaching model, which included pre-class human–computer interactive oral training, in-class in-depth human–computer dialogue and PPT presentations, and after-class human–computer interactive PPT video production. The control class (30 students, 2025 intake, Civil Engineering) followed the conventional oral English teaching approach. The research spanned one semester, and data were collected through final oral examination performance analysis and face-to-face interviews. The results indicate that students in the experimental class significantly outperformed the control class in terms of oral performance with the mean final oral score exceeding that of the control class by 1.47 points ($p = 0.002$). Meanwhile, AI-assisted oral training also exerted a positive transfer effect on students' comprehensive language proficiency (reading, translation, and writing). This study indicates that AI technology can positively impact L2 oral instruction, suggesting instructors consider revising L2 speaking pedagogy. In addition, the experiment revealed that over-reliance on AI may weaken students' autonomous thinking ability, so targeted suggestions for mitigating such risks are accordingly proposed.

Keywords: *AI-enabled instruction; second language oral proficiency; human–machine interaction; instructional reform*

This is an open access article under the CC BY-NC license.

DOI: 10.5281/zenodo.24126238

1. Research Background

Against the backdrop of economic globalization and increasingly frequent cross-cultural communication, China's requirements for college students' English oral proficiency have continued to rise. Fluent, accurate, and appropriate oral expression has become one of the core competencies for college graduates in their professional work and international engagement. However, students cultivated through traditional instructional models often exhibit high test scores but low communicative competence; even those who pass the high-level English proficiency exams may still struggle to communicate effectively in authentic situations. Consequently, leveraging artificial intelligence (AI) to break through the limitations of traditional instruction and to construct an efficient, feasible oral English teaching model is not only a practical necessity for improving teaching quality, but also a strategic choice for cultivating versatile talents that meet the demands of the times.

1.1 Common Problems in L2 Instruction in Chinese Universities

College English, as a core public course in Chinese higher education, bears the important responsibility of cultivating students' cross-cultural communicative competence and language application ability. Currently, L2 oral instruction in Chinese universities faces three primary challenges. First, class sizes are large, affording students few opportunities to speak and delaying teacher feedback. In China, it is common for a College English class to contain 80–100 students; such large-scale instruction results in limited speaking opportunities and seriously constrains the instructor's understanding of students' actual performance while also delaying feedback. Pronunciation errors and inappropriate expressions in oral practice are difficult to correct in a timely manner, and students often repeat errors without being aware of them. Second, instruction often lacks authentic communicative contexts; role-play activities are limited, lightly monitored, and largely confined to everyday topics – falling far short of the needed training in higher-order oral scenarios such as academic discussion and cross-cultural debate. The instructional model is rigid and monolithic, relying excessively on textbooks and classroom lecturing, lacking immersive and interactive oral practice experiences, and thus failing to engage students or activate their linguistic competence. Third, students tend to avoid expression due to speaking anxiety (Cheng, 2017). Foreign language speaking anxiety is a prevalent psychological phenomenon with multidimensional negative effects (Rahman & Tomy,2024). Offering interaction opportunities with an AI can significantly lower anxiety.

1.2 The Opportunity of AI-Enabled Education

Relying on core technologies such as big data, natural language processing (NLP), automatic speech recognition (ASR), and text-to-speech synthesis (TTS), AI provides strong support for the reform of college English oral instruction (Zhang, 2019; Pang, He & Marlow, 2025; Pang, Marlow & He, 2025). Specifically, AI has demonstrated prominent application value in oral interactive training (Kong, 2024; Pang & Marlow, 2025b) and AI-based speaking assessment (Jeon & Lee, 2023). Large language models (LLMs) such as ChatGPT and *Doubao* possess multimodal interaction, contextual understanding, and real-time content generation capabilities, enabling them to simulate authentic communicative situations, precisely identify learners' linguistic errors, and provide immediate feedback. Furthermore, AI technology helps alleviate learners' speaking anxiety (Dizon & Tang, 2020; Chen et al., 2022), increases the frequency of oral output, and enhances the accuracy of language expression. Wu et al. (2025) further found that embedding chatbots into cooperative learning models can create a low-pressure oral communication environment, enhance learners' sense of learning enjoyment, and simultaneously reduce speaking anxiety.

2. Relevant Theoretical Framework

While AI is new, learner challenges with input, output, and context are ongoing.

2.1 The Input and Output Hypotheses in SLA

Krashen's (1985) Input Hypothesis emphasizes that comprehensible input is the key to language acquisition. Its "i+1" principle holds that language input should be slightly above the learner's current level, ensuring comprehensibility while promoting competence development through appropriate challenge. In today's oral instruction, AI technology can generate topic-appropriate materials and model expressions tailored to students' proficiency levels, precisely controlling lexical difficulty and syntactic complexity, thereby achieving personalized "i+1" input and laying the foundation for subsequent oral output with individualization not possible for a single instructor in a large classroom.

Swain's (1995) Output Hypothesis argues that language acquisition depends not only on sufficient input but also on meaningful language output. Output activities such as oral expression prompt learners to actively organize linguistic knowledge, identify their own deficiencies, and thereby achieve language internalization. AI technology provides efficient support for oral output training: intelligent chatbots and virtual language partners can offer round-the-clock, pressure-free oral interaction scenarios, providing real-time feedback on pronunciation, grammar, and expression problems, and encouraging students to practice and revise repeatedly.

2.2 Situated Learning Theory

Situated Learning theory (Lave & Wenger, 1991) holds that learning is a process in which learners interact with the environment and construct knowledge within authentic situations. The development of oral competence depends particularly on authentic communicative contexts, yet traditional instruction struggles to reproduce diverse oral usage situations. Through virtual situation simulation, AI technology can construct immersive oral scenarios such as cross-cultural negotiation, academic seminars, and everyday debates, enabling students to practice fluency and appropriateness through role-playing. Group collaboration can also be incorporated to combine AI-generated tasks and realize integrated online–offline situated oral practice, helping students transform linguistic knowledge into practical communicative competence.

2.3 Data-Driven Learning Theory

Data-Driven Learning (Johns, 1991) centers on massive learning data and provides precise pedagogical support through data analysis and mining. In the Chinese context, the intelligent agents on the *Chaoxing* (Gao, et al, 2026) learning management system (LMS) can collect behavioral data such as students' oral interaction frequency, types of pronunciation errors, topic preferences, and assignment completion quality. Through big data analysis, these agents can accurately identify students' weaknesses and push

This is an open access article under the CC BY-NC license.
DOI: 10.5281/zenodo.24126238

personalized oral training resources; teachers can then adjust teaching priorities based on data feedback, achieving teaching guided by learning and enhancing the pertinence and effectiveness of oral instruction. Although the present study employed *Chaoxing*, the pedagogical framework should apply to other LMS environments offering comparable functionality.

2.4 Applying Theory to Practice

The above theories provide theoretical support for AI-enabled L2 oral instruction from different dimensions and align explicitly with the four practical problems described in section 1.1. Specifically, the Input and Output Hypotheses, through AI-generated *i*+1 materials and round-the-clock, one-on-one output opportunities with immediate feedback, jointly address limited speaking opportunities and delayed feedback. Situated Learning theory, by constructing immersive communicative contexts through AI virtual simulation, directly connects to the lack of authentic contexts; and all three theories jointly contribute to anxiety alleviation—the non-judgmental, and non-human nature of AI language partners provides a low-risk space for trial and error, while Data-Driven Learning precisely identifies weaknesses and pushes personalized resources, keeping practice within the learner's zone of proximal development (Vygotsky, 1978) and preventing frustration and aggravation driven by anxiety. Together, the three theories constitute a complementary framework at the levels of acquisition functions, learning context, and pedagogical decision-making, laying a theoretical foundation for minimizing these difficulties.

3. Experimental Study

3.1 Participants

This study selected two parallel classes of non-English-major students from related majors at the author's (regional, comprehensive, public) university as participants. Regional undergraduate institutions account for a high proportion of Chinese universities, and their students typically possess moderate English proficiency; the previously identified problems of limited speaking opportunities, delayed feedback, lack of authentic contexts, and high speaking anxiety in their oral instruction are widespread. Moreover, these challenges are not limited to China alone but also apply in the United States and elsewhere. The current research was conducted over the first semester of the 2025–2026 academic year, lasting 16 weeks, with one 40-minute oral lesson per week. The experimental class comprised 28 students, and the control class had 30 students. At the start of the semester, all students took an English proficiency entry test (comprising six components: oral, listening, grammar, reading, translation, and writing, with a total score of 100). The experimental class had a mean score of 68.3 and the control class 67.9; an independent-samples t-test yielded $t = 0.523$, $p > 0.05$.

3.2 Experimental Variables

Our independent variable is instructional model. The experimental class adopted a human–machine interactive oral teaching model; the control class followed a traditional oral teaching model (textbook dialogue practice, group discussion, and teacher oral commentary).

Several dependent variables are assessed: students' English oral proficiency (measured primarily by the oral component score on the final examination), learning interest, and comprehensive language proficiency (reading, translation, and writing scores).

Control variables include the textbook (New College English 1 Integrated Course, 4th edition, Foreign Language Teaching and Research Press), instructor, class schedule, and difficulty of the oral examination questions.

3.3 Implementation of the Experimental Class

This study relies on the intelligent agents of the *Chaoxing* language learning platform, developed and operated by Beijing Century Superstar Information Technology Development Co., Ltd., which integrates dialogue simulation, speech assessment, data collection, and resource recommendation within a single system, fitting the practical needs of this research. Accordingly, a full-process oral teaching model was constructed: AI-enabled input — human–machine interactive output — practical application and reinforcement — intelligent evaluation and feedback. This model comprises the following modules.

First, pre-class: AI-driven personalized oral input and rehearsal. Based on students' interests and proficiency levels, the instructor used large language models (e.g., *Doubao*) to generate diverse oral topic materials (e.g., social hotspots, cultural differences, technology ethics), accompanied by model expressions, key vocabulary, and sentence frameworks. Students received preview tasks through the *Chaoxing* platform and engaged in topic rehearsal dialogues with the AI language partner, which corrected pronunciation and grammar errors in real time and provided suggestions for expression improvement. Through dialogue with the intelligent agent, students accumulated language resources related to the topics.

Second, in-class: human–machine collaborative situated interaction and cooperation. A dual-track model of "human–machine interaction + interpersonal collaboration" was adopted. First, students in groups conducted situational simulated dialogues with the AI language partner on the *Chaoxing* platform or with self-developed intelligent agents. The AI played different roles (e.g., debate opponent, customer, interviewer), driving students to further elaborate their viewpoints on previously prepared topics through follow-up questioning, challenging, and rebuttal, thereby honing logical thinking and adaptability. Subsequently, each group produced PPT presentations based on the dialogue outcomes and delivered in-class presentations; the AI detected in real time the accuracy, syntactic diversity, and coherence of the language in the PPT and provided revision suggestions. Finally, the teacher organized whole-class commentary focusing on the authenticity, logical coherence, and communicative appropriateness of oral expression, helping students develop critical thinking and deep reflection.

Third, post-class: AI-assisted video production and multidimensional evaluation. Students produced videos on the topics discussed before and during class. AI tools supported script optimization (supplementing cultural details and polishing expressions), speech synthesis (adjusting speech rate and accent), and subtitle generation. After uploading the videos to the *Chaoxing* platform, channels for peer review and teacher commentary were opened, encouraging students to leave comments from three dimensions—linguistic accuracy, expressive fluency, and content completeness—thereby promoting reflection and improvement.

3.4 Implementation of the Control Class

The control class followed a traditional oral teaching model, forming a contrast with the experimental class. Oral instruction

This is an open access article under the CC BY-NC license.
DOI: 10.5281/zenodo.24126238

centered on textbook dialogue reading aloud, group topic discussion, and role-playing. The teacher explained pronunciation and grammar rules; after students' oral presentations, the teacher provided holistic commentary without individualized feedback. Post-class assignments consisted of written work (e.g., dialogue script writing, topic essays) that the teacher graded manually, with a feedback cycle of approximately 3–5 days; no video production or AI-assisted optimization was involved.

4. Discussion and Analysis

4.1 Comparison of Final Oral Examination Scores

The oral component of the final English examination was scored out of 15. The results of the two groups are presented below (Table 1).

Table 1. Comparison of oral proficiency scores between the experimental and control classes

Class	N	Oral Mean Score (out of 15)	SD
Experimental	28	12.50	2.12
Control	30	11.03	0.81

The mean oral score of the experimental class (M = 12.5, SD = 2.12) was higher than that of the control class (M = 11.0, SD = 0.81); an independent-samples t-test indicated that the difference was statistically significant, $t(34.3) = 3.44$, $p = 0.002$.

4.2 Comparison of Final Comprehensive Language Proficiency Scores

The comprehensive language proficiency component of the final examination comprised reading, translation, and writing, scored out of 55. The results of the two groups are presented below (Table 2).

Table 2. Comparison of comprehensive language proficiency scores between the experimental and control classes

Class	Reading Mean (of 30)	Translation Mean (of 10)	Writing Mean (of 15)	Comprehensive Mean (of 55)	SD
Experimental	24	8	12	44	4.68
Control	22	8	11	41	2.49

The comprehensive mean score of the experimental class (M = 44.00, SD = 4.68) was 3 points higher than that of the control class (M = 41.00, SD = 2.49); an independent-samples t-test indicated that the difference was statistically significant, $t(40.6) = 3.02$, $p = 0.004$. This suggests that AI-driven listening and speaking instruction may exert a positive effect on reading, translation, and writing abilities. One possible explanation is that the AI-supported activities provided additional lexical, grammatical, and written language practice that may have contributed to performance in these areas.

4.3 Analysis of Face-to-Face Interview Results

Following the experiment, face-to-face interviews were conducted with 20 students from the experimental class and 15 students from the control class. Combined with teaching logs, the advantages and problems of AI-enabled oral instruction were identified as follows.

4.3.1 Notable Advantages

Enhanced interest and initiative in oral learning: 85% of experimental-class students reported that AI-generated real-time topics and situational tasks were more engaging than textbook content, being “close to real life” and “challenging,” thereby stimulating their willingness to practice actively; in contrast, 60% of control-class students reported that the topics were “outdated” and “lacking novelty,” with low levels of motivation.

Increased expressive confidence and reduced anxiety: 78% of experimental-class students indicated that the pressure-free interaction environment provided by the AI language partner enabled them to “dare to speak without fear of making mistakes,” and that the AI's immediate encouragement and suggestions enhanced their self-efficacy; control-class students generally reported that “oral practice received no timely feedback, and they did not know whether what they said was correct.”

Improved learning efficiency: experimental-class students reported that AI's immediate revision and personalized recommendations saved considerable time otherwise spent consulting reference materials and waiting for feedback, enabling them to quickly identify weaknesses and make targeted improvements; the teacher likewise noted that AI-assisted learning analytics reduced repetitive workload, allowing them to focus more on cultivating higher-order oral competence.

4.3.2 Identified Problems and Suggestions for Improvement

Over-reliance on AI: 70% of experimental-class students admitted that, when writing scripts and optimizing PPT copy, they directly adopted the AI's revision suggestions, some reported being “too lazy to work out ways of expression on their own”; some students' self-reported first reaction when encountering expressive difficulties was to resort to AI rather than attempt independent problem-solving, resulting in a decline in autonomous problem-solving ability.

Insufficient technical adaptability: students also reported concerns over AI's recognition accuracy, suggesting that the AI language partner had limited depth of comprehension in complex topic discussions, “being unable to respond to some abstract or speculative content.” In terms of accent training, students also doubted the feedback reporting repeated instances of “correct pronunciation yet incorrect AI judgment” for unique situations.

Optimize instructional design to balance AI use with autonomous thinking: reduce AI functions that directly provide answers, and increase open-ended and inquiry-based tasks, such as designing oral activities in which “the AI poses questions and students independently collect materials and construct arguments”; require students to annotate the reasons for their revisions after using AI, so as to reinforce deep understanding of linguistic rules.

Strengthen teacher guidance to leverage the advantages of human–machine collaboration: teachers should, based on the learning-analytics data of their LMS platform, precisely identify students' dependent behaviors and design targeted word-prediction activities.

This is an open access article under the CC BY-NC license.
DOI: 10.5281/zenodo.24126238

Improve technical adaptability to enhance the learning experience: it is recommended that the LMS platforms optimize their AI models to improve recognition accuracy for non-standard accents and comprehension of complex contexts; and introduce functions that limit AI usage duration and prompt autonomous thinking (e.g., a reminder to try to infer from context after three consecutive AI word-lookup queries), so as to guide students in the rational use of the tool.

5. Conclusion

Through a quasi-experimental design, this study verified the feasibility and effectiveness of an AI-assisted college English oral teaching model based on the intelligent LMS agents in improving students' oral proficiency. The results show that the experimental class significantly outperformed the control class not only in oral accuracy and fluency but also in comprehensive language proficiency. By leveraging AI-generated personalized topic materials and human–machine interactive training, the model effectively responded to the practical difficulties of traditional oral classrooms, including insufficient willingness to speak, delayed feedback, and a lack of authentic communicative contexts; moreover, the intelligent evaluation system embedded in the LMS platform provided learners with immediate feedback and precise learning guidance. Meanwhile, the study also identified problems in practice, such as learners' over-reliance on AI, which can be mitigated through optimizing instructional design, strengthening the teacher's scaffolding role, and improving technical adaptability. The instructional model constructed in this study is not only applicable to the development of oral English proficiency but also offers referential value for L2 oral instruction in other languages.

The study has two core limitations: a relatively small sample size and an intervention period of only one semester. Future research could expand the sample size and extend the experimental and follow-up periods, examining the long-term retention of L2 oral proficiency, so as to conduct a more comprehensive evaluation of the sustainability and language transfer effects of the AI-enabled oral teaching model.

ACKNOWLEDGEMENT

This paper is supported by the 2025 Guangdong Provincial Education Science Planning Project (Scientific Research Topics on Higher Education), entitled *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. Chen, C., Koong, C., & Liao, C. (2022). Influences of integrating dynamic assessment into a speech recognition learning design. *Educational Technology & Society*, 25, 1-14.
2. Cheng, Y. S. (2017). Development and preliminary validation of four brief measures of L2 language-skill-specific anxiety. *System*, 68, 15–25. https://doi.org/10.1016/j.system.2017.06.009
3. Dizon, G., & Tang, D. (2020). Intelligent personal assistants for autonomous second language learning. *JALT CALL Journal*, 16, 107-120.
4. Gao, Y., Teo, T., Wong, S. L., Geng, J., & Diao, F. (2026). Unpacking Chinese university EFL teachers' continuance intention toward LMS: A qualitative extension of the UTAUT2 framework. *SAGE Open*, 16(3).
5. Jeon, J., & 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.
6. Johns, T. (1991). Should you be persuaded: Two examples of data-driven learning. *English Language Research Journal*, 4, 1-16.
7. Kong, L. (2024).Application of GAI in foreign language teaching: A case study. *Foreign Language Education in China*, 7(1), 11-18, 90.
8. Krashen, S. D. (1985). *The input hypothesis: Issues and implications*. Longman.
9. Lave, J. & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge University Press.
10. Pang, Y.F.; He, Y.Y. & Marlow, D. (2025). Employing large language models in college English teaching: Insights from China. *International Journal of Social Sciences, Language and Linguistics*, 5(6), 16-21. https://doi.org/10.55640/ijsll-05-06-03
11. Pang, Y.F. & Marlow, D. (2025). Application of artificial intelligence in college English listening teaching. *International Journal of Social Sciences, Language and Linguistics*, 10, 18-24.
12. Pang, Y.F.; Marlow, D.& He, Y. (2025). Research on empowering college English teaching with artificial intelligence. *Education and Social Work*, 1(1). https://doi.org/10.63313/ESW.9015
13. Rahman, A., & Tomy, P. (2024). Intelligent personal assistant: An interlocutor to mollify foreign language speaking anxiety. *Interactive Learning Environments*, 32(8), 4569-4586. https://doi.org/10.1080/10494820.2023.2204324
14. Swain, M. (1995). *Three functions of output in second language learning*. In G. Cook & B. Seidlhofer (Eds.), *Principle and practice in applied linguistics: Studies in honor of H. G. Widdowson* (pp. 125-144). Oxford University Press.
15. Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes* (M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, Eds.). Harvard University Press.
16. Wu, T.T.; Hapsari, I. P. & Huang, Y.M. (2025). Effects of incorporating AI chatbots into think-pair-share activities on EFL speaking anxiety, language enjoyment, and speaking performance. *Computer Assisted Language Learning.* https://doi.org/10.1080/09588221.2025.2478271

This is an open access article under the CC BY-NC license.
DOI: 10.5281/zenodo.24126238

17. Zhang, Y.L. (2019). Research on the Transformation of College English Teaching Driven by Artificial Intelligence. *Education Modernization*, 6(55), 130-131.

This is an open access article under the CC BY-NC license.
DOI: 10.5281/zenodo.24126238