

Employing Large Language Models in College English Teaching: Insights from China

Yafei Pang^{1*}, Yuyan He², David Marlow³

¹Foshan University, Foshan, Guangdong, China

²Foshan University, Foshan, Guangdong, China

³University of South Carolina Upstate, Spartanburg, South Carolina, USA

Doi <https://doi.org/10.55640/ijssl-05-06-03>

ABSTRACT

This study focuses on the innovative application of large language models (LLMs) in college spoken English teaching. Based on the ever-evolving principles of Human-Computer Interaction (HCI), it systematically analyzes the functional positioning and implementation paths of LLMs in three traditionally core educational roles: language consultant, conversational language partner, and oral production assessor. Through case studies of prompt engineering in teaching scenarios such as listening and speaking training, situational dialogue, and real-time feedback, this paper explores the practical value of LLMs in pronunciation correction, academic oral training, and cross-cultural communication simulation. The research shows that while LLMs still face multiple challenges at the technical, pedagogical, and ethical levels, they can effectively compensate for the traditional shortcomings of oral language instruction at the college level, such as insufficient interaction and delayed feedback.

Keywords: Large Language Models (LLMs), Human-Computer Interaction (HCI), Prompt Engineering, College English Teaching, Insights from China.

INTRODUCTION

University-level teaching of spoken English as an international language has long faced three major challenges: limited opportunities to speak (Savaşçı, 2014; Shumin, 2002; Suban, 2021), delayed feedback (Hossain, et al, 2004; Jia & Lu, 2025; Rassaei, 2023), and inauthenticity of scenarios (Barekat & Nobakhti, 2014). In traditional classrooms, teachers struggle to provide personalized guidance for classes of fifty students or more (Hai-yan, 2014; but see Jin & Cortazzi, 1998 for an argument for large class sizes). Students often avoid speaking aloud due to anxiety (Liu, 2005 & 2006), and role-playing exercises are generally limited to everyday topics, lacking authentically advanced scenarios such as academic discussions and cross-cultural interaction (Lin, 2009). Large Language Models (LLMs), a subset of Artificial Intelligence tools focused on communication like ChatGPT and Wenxin Yiyuan, offer new possibilities to break through these bottlenecks with their multimodal interaction, contextual understanding, and real-time response capabilities (Wei, et al, 2025). Their automated speech recognition (ASR) (Liu, et al, 2022) and natural language processing (NLP) (Ou, 2023) technologies can simulate real dialogue scenarios, identify

language errors in real time, and provide precise feedback, making them important tools for improving students' oral fluency and pragmatic competence.

1.1 Research Status and Innovative Perspective

Early studies in Human-Computer Interaction (HCI) in education have verified the positive impact of intelligent language partners on oral learning: Dizon & Tang (2020) found that conversing with intelligent assistants like Alexa reduced students' anxiety by 32% and increased speaking frequency by 41%; Chen et al. (2022) proved through experiments combining dynamic evaluation and speech recognition that technological intervention can improve oral accuracy by 28%. Research within China as focused more on tool function discussions. For example, Xu Jiajin et al. (2024) proposed that LLMs can play the roles of *language consultant*, *language partner*, and *assessment expert* in English teaching, and Qin, et al (2025) demonstrated a 13% decrease in failure rates with HCI enhanced instruction. Still, in-depth applied research on oral teaching, especially in core college English scenarios such as academic oral training and cross-cultural

communication simulation, still lacks systematic exploration—the specific paths and effects of educationally focused HCI need further discussion.

1.2 Research Objectives and Methods

This study has two key aims in support of the teaching of oral English. First, to construct a role framework for LLMs and, second, to analyze specific strategies for technological application of prompt engineering. The research uses a literature analysis method to sort out theoretical foundations and then builds on these foundations to explore potential roles and prompts for typical scenarios from teaching practice to assist teachers as they seek operational technology integration for their students.

2. Core Teaching Roles and Functions of Large Language Models

2.1 Language Consultant: Precise Guidance on Pronunciation and Grammar

Used as an “intelligent language coach”, LLMs can analyze speech, vocabulary, and grammar issues in students' oral output and provide professional guidance in real-time.

2.1.1 Pronunciation Correction and Intonation Simulation

The sound system for each language creates unique issues for learners' pronunciation. In Chinese students, for example, common problems include confusion between /l/ and /r/ and distinguishing between long and short vowels. LLMs can analyze acoustic features of students' speech and generate visual feedback. For example, after a student reads the word “world”, the model might feedback: “Your pronunciation of /ɜ:/ is closer to /ə/, which may cause misunderstanding. Listen to the audio sample and repeat: [audio link]. The provided spectrogram may help you see the difference”. This technical enhancement to the traditional “shadowing method” where students listen and repeat native speakers' speech. Not only does the enhanced model compare students' speech with that of native speakers' word-by-word or sentence by sentence, marking intonation differences (e.g., rising tones at the end of declarative sentences) but it provides immediate, unambiguous, and individualized feedback.

2.1.2 Academic Vocabulary and Pragmatic Knowledge Support

In academic oral scenarios, the model can generate collocation networks for relevant terms based on dialogue topics, enabling students to choose topics suited to their own interests. For example, for a student wanting to explore AI ethics, the instructor could enter the prompt: “Act as a

computer science professor. Suggest academic vocabulary for discussing AI ethics, such as ‘algorithmic bias’ and ‘data privacy’, and explain their contextual usage”. The model will then return terms like “surveillance capitalism” and “human-centric design”, providing definitions and example sentences: “When debating AI ethics, it is crucial to address how surveillance capitalism intersects with data privacy concerns”. This topical individualization enables authentic learning that is not possible in the traditional classroom.

2.2 Intelligent Language Partner: Immersive Communication Scene Construction

Building on vocabulary and knowledge supports, it is time to explore how the LLMs can break through traditional time and space limitations to simulate individualized dialogue partners to create low-pressure practice environments and authentic scenarios with a high degree of cognitive challenge.

2.2.1 Academic Discussion and Critical Thinking Training

In advanced oral teaching, the model can function as an academic seminar participant to guide students in argumentation and rebuttal (Zhang, 2023). For example, in a discussion on the topic “*The Impact of Large Language Models on Language Learning*”, the model might pose a question as a critic: “Critics argue that relying on LLMs for language practice reduces human interaction. How would you defend the value of AI companions in academic discourse?” Students would need to use logical reasoning (e.g., citing Edmett et al.'s 2023 research on *technology supplementing human communication*) to respond, thereby authentically employing critical thinking skills.

2.2.2 Daily Dialogue and Cross-Cultural Communication Simulation

In basic oral training, the model can play roles such as an international roommate or tour guide, engaging in interactions around campus life, travel planning, and other topics (Li & Lu, 2024). For example, setting the prompt: “You are a Canadian exchange student living in Beijing. Talk with the student about cultural differences in dining habits, and ask follow-up questions to deepen the conversation”, the model might actively ask: “In China, I notice many people share dishes at restaurants. How does this differ from typical meals in your hometown?” Such interactions help students build response strategies in real-life situations.

2.3 Oral Assessment Expert: Multi-Dimensional Competency Evaluation

Traditional oral tests rely on subjective teacher scoring, suffering from vague criteria and delayed feedback. LLMs, combined with ASR and NLP technologies, can build an instant diagnosis and dynamic tracking assessment system (Zheng, et al, 2025).

2.3.1 Quantitative Analysis of Fluency and Coherence

The model can automatically calculate indicators such as speaking rate (words per minute), filled pause frequency, and topic continuation length. For example, after a student completes a 2-minute presentation, the model might generate a report: *"Speaking rate: 105 wpm (below the academic benchmark of 120 wpm); filled pauses (um/uh): 4 times per minute; average turn length: 3.5 sentences – indicating limited elaboration."* Such data helps instructors understand students' needs and helps students make targeted improvements.

2.3.2 Pragmatic Appropriateness Assessment

In cross-cultural communication scenarios, the model can evaluate the appropriateness of language use in social situations. For example, when a student simulates a dialogue to *"request an extension from a professor"*, the model might point out: *"Your request 'I need more time, so just extend the deadline' is too direct. In academic contexts, it's more appropriate to say 'Would it be possible to request a one-week extension due to unexpected circumstances?'"* and explain the tone differences between formal and informal contexts.

3. Prompt Engineering in Teaching Scenarios

With HCI-enhanced teaching, prompt engineering can effectively improve instructional relevance and students' language development by optimizing model interaction instructions (Hegazy, 2024; Roy 2024). In language teaching practice, its applications can be divided into three typical scenarios: listening and discrimination training, situational dialogue, and real-time feedback, each achieving specific teaching objectives through differentiated process designs.

3.1 Prompt Engineering: Listening and Discrimination Training

This scenario aims to systematically correct grammatical and lexical errors in students' oral speech through immediate intervention. Its prompt engineering process is built on intelligent speech analysis technology: first, the model uses natural language processing algorithms to real-time monitor students' speech input, accurately identifying grammatical issues such as subject-verb agreement errors and tense confusion, as well as misuse of lexical collocations; in feedback strategies, it combines explicit correction with guided revision. Explicit correction directly points out errors and

provides standard expressions—for example, for the sentence *"She go to school yesterday"*, the feedback is: *"In your sentence, 'She go to school yesterday', the past tense of 'go' is 'went'. Correct sentence: 'She went to school yesterday'."* Guided revision, on the other hand, prompts error points to inspire students to think independently about revisions—for example, for *"I very like this book"*, the prompt is: *"You said 'I very like this book'. In English, 'very' does not modify verbs. How might you rephrase this?"* This cultivates students' autonomous error-correction abilities. In most situations, the model should be directed to give students guided revision first and then to provide explicit correction upon the student's request.

3.2 Prompt Engineering: Situational Dialogue

Situational dialogue in college English teaching is an interactive training model designed around specific language application scenarios, aiming to systematically improve students' oral comprehension and expression abilities through simulated real communication contexts. This training model follows a structured framework of task decomposition (also known as prompt guidance), especially suitable for scenarios requiring precise language output, such as practicing for an academic defense.

To use LLMs to aid in training students for academic defense responses, prompt engineering best practice includes two core components: task decomposition and prompt design. To break apart the primary defense goal of clearly responding to judges' questions the LLMs may define three progressive steps: identifying the core of the question, organizing the response structure, and oral expression of the response. In the prompt design stage, guided instructions are generated according to the specific scenario of the student's defense. For example, for the common question of research innovations, the LLMs can design a task: *"Suppose a judge asks, 'What are the breakthroughs of your research compared to previous studies?' You should first identify the key phrases in the question, 'research breakthroughs' and 'comparison with predecessors', then organize your answer using the following structure: first explain the difference in research perspective, second list specific methodological improvements, and finally summarize the innovative value.' After you respond, I will simulate the judge's further questioning."* As needed, basic response templates may be provided for students to reference, such as: *"The innovations in my research are mainly reflected in two aspects: first, in methodology... (methodological innovation), second, in conclusions... (conclusion breakthroughs), which address the limitations in previous research..."* This training method of transforming complex

language tasks into actionable steps helps students gradually master the core skills of grasping question cores as they build logical frameworks and work toward fluent oral expression in realistic defense scenarios, shifting from generic passive learning to authentic and active language production.

3.3 Prompt Engineering: Real-Time Feedback

Real-time feedback is a teaching model of immediate evaluation and guidance in language interaction, helping learners optimize expression strategies through the closed loop of scenario simulation and immediate response. In cross-cultural communication training, this model enables students to intuitively understand communication differences across cultures.

Take “simulating German business negotiations” as an example: first, the LLMs sets roles, with the model playing a German procurement manager who values data-based logic, and the student playing a representative of a Chinese supplier. The ensuing dialogue may then focus on product price increases. Prompt design incorporates the characteristics of German culture that emphasize directness and precision, giving the task: *“When the German side says, ‘Your price increase is too sudden, we need justification,’ please first clarify that the other party’s core concern is a logical basis for the price, then respond in two steps: 1) explain cost changes using the percentage increase in raw materials, and 2) compare the quality differences of competing products.”* In the interaction, if the student responds with something generic like, *“Market prices are all increasing”*, the model will simulate the German side’s follow-up question: *“How much has your raw material cost increased? Please show a price comparison with last year.”* Further, the model can provide real-time feedback: *“Your response lacks specific data—Germans are accustomed to supporting arguments with numbers. You should add quantitative information such as ‘a 25% increase in steel prices this year has led to an 18% increase in production costs’ to make the reasoning more persuasive.”* This targeted training method enables international business students to experience more authentically cross-cultural communication challenges than could ever happen in a traditional classroom, enhancing their adaptive expression capabilities in realistic scenarios.

4. Human-Computer Interaction: Advantages and Challenges

4.1 Core Advantages

HCI innovates language learning models through intelligent technology, demonstrating significant advantages in personalized teaching and authentic scenario expansion. LLMs can dynamically adjust interaction difficulty according to students’ language proficiency, designing ladder

questions for those with weaker foundations and creating argumentation scenarios for advanced learners to achieve customized practice tailored to their needs. Meanwhile, anonymous interaction mechanisms significantly reduce students’ fear of evaluation (Demir, et al, 2024); research shows that students using this model have a 54% higher willingness to speak in class (Xu, 2004; see also Peng & Liang, 2005). Additionally, the technology breaks through the individuality and authenticity limits on scenarios in traditional teaching environments, simulating unique training scenarios such as academic defenses and cross-cultural negotiations without the need for real language partners, providing multiple paths for language ability advancement.

4.2 Realistic Challenges

Current HCI-enabled teaching still faces multiple challenges at the technical, pedagogical, and ethical levels. Technically, the system has accent recognition biases, with an accuracy rate of only about 78% for non-standard English variants (such as Indian English, African English) and Chinese dialects, and it cannot capture non-verbal communication signals such as eye contact and gestures. In terms of pedagogical adaptation, instructors may make the unfortunate assumption that technology can replace teaching, leaving students to interact with LLMs on their own (Chan & Tsi, 2023). Similarly avoiding personal responsibility, students who choose to can easily use an LLMs of their own to imitate their personal effort (Shin, et al, 2025), foregoing the critical thinking and learning that HCI-enabled teaching can foster. Another significant challenge lies in the realm of ethics (Kumari, et al, 2025). Voice data storage may lead to privacy leakage risks, and the model’s lack of emotional interaction capabilities may hinder the development of students’ social and emotional skills due to long-term reliance.

5. Conclusions and Prospects

Large Language Models have brought a paradigm shift to college spoken English teaching: their threefold roles as language consultant, intelligent language partner, and assessment expert effectively solve the problems of insufficient interaction, delayed feedback, and single scenarios in traditional teaching, particularly demonstrating unique value in individualized academic oral training and cross-cultural communication simulation. However, when applying this or any technology, we must guard against the risk of dehumanization, always maintaining the improvement of students’ real-world communication abilities as the core goal. In the era of generative artificial intelligence, college

English teachers should actively embrace technology, transforming from knowledge controllers to educational HCI designers, allowing LLMs to become a pivot for improving oral teaching innovation and helping students develop authentic language abilities in individualized virtual scenarios.

Looking to the future, with the continuous progress and improvement of AI technology, its applications in college English teaching will become more in-depth and extensive. This will not replace human teachers but will enable college English teaching to reach greater heights. By using AI to orchestrate repetitive linguistic fundamentals and learner-specific advanced situational training, it will be possible to cultivate high-quality English-proficient talents who meet the authentic real-world communicative needs of our increasingly globalized world, providing a strong talent pool for international exchanges and collaborative initiatives.

Acknowledgments: This article is a phased achievement of the University-level AI Curriculum Project "Research on the Comprehensive Enhancement of College English Listening and Speaking Abilities Driven by AI" (No. PX-98251466) by Foshan University.

REFERENCES

1. Barekat, B., & Nobakhti, H. (2014). The effect of authentic and inauthentic materials in cultural awareness training on EFL learners' listening comprehension ability. *Theory and Practice in language studies*, 4(5), 1058.
2. Chan, C.K.Y.; Tsi, L.H. 2023. The AI Revolution in Education: Will AI Replace or Assist Teachers in Higher Education? arXiv:2305.01185.
3. Chen, C., Koong, C., & Liao, C. (2022). Influences of integrating dynamic assessment into a speech recognition learning design. *Educational Technology & Society*, 25, 1-14.
4. Demir, O., Keskin, S., & Cinar, M. (2024). The effect of social anxiety on student interactions in asynchronous online discussion forums as mediated by social presence and moderated by anonymity. *Journal of Computing in Higher Education*, 1-30.
5. Dizon, G., & Tang, D. (2020). Intelligent personal assistants for autonomous second language learning. *JALT CALL Journal*, 16, 107-120.
6. Edmett, A., et al. (2023). *Artificial Intelligence and English Language Teaching*. London: British Council.
7. Hai-yan, M. (2014). The Task-Based Teaching of Writing to Big Classes in Chinese EFL Setting. *English Language Teaching*, 7(3), 63-70.
8. Hegazy, H. M. (2024). A Linguistically Developed Prompt Engineering Parameters Model for Enhancing AI's Generation of Customized ESL Reading Texts* Dr. Hebatollah MM Hegazy. *Faculty of Education Journal Alexandria University*, 34(3), 501-566.
9. Hossain, K. I., Ahmed, M. K., & Mahmud, M. S. (2024). A Comprehensive Review on the Impact of Teacher Feedback in Language Development for ESL/EFL Learners. *IUBAT Review*, 7(1), 218-229.
10. Jia, S., & Lu, Z. (2025). Enhancing EFL Oral Production through Mobile-Assisted Task-Based Language Teaching: A Study in Effectiveness. *Journal of Information Technology Education: Research*, 24, 012.
11. Jin, L., & Cortazzi, M. (1998). Dimensions of dialogue: Large classes in China. *International journal of educational research*, 29(8), 739-761.
12. Kumari, P., Swarnalatha, K. S., Naveen, G., Sale, A., Prajwal, M. P., & Siddhardha, K. S. (2025). Challenges and Ethical Considerations of AI in English as a Foreign Language. In *Application of AI in the Teaching and Learning of English as a Foreign Language (EFL)* (pp. 149-178). IGI Global Scientific Publishing.
13. Li, C., & Lu, L. (2024, November). AI-Enhanced Intercultural Teaching: Investigating Chinese EFL University Teachers' Technology Readiness and Acceptance of Large Language Models. In *International Symposium on Emerging Technologies for Education* (pp. 123-134). Singapore: Springer Nature Singapore.
14. Lin, Y. E. (2009). Investigating role-play implementation: A multiple case study on Chinese EFL teachers using role-play in their secondary classrooms.
15. Liu, J., Liu, X., & Yang, C. (2022). A study of college students' perceptions of utilizing automatic speech recognition technology to assist English oral proficiency. *Frontiers in Psychology*, 13, 1049139.
16. Liu, M. (2005). Reticence in Oral English Language Classrooms: A Case Study in China. *TESL Reporter*, 38(1), 1-16.
17. Liu, M. (2006). Anxiety in Chinese EFL students at different proficiency levels. *System*, 34(3), 301-316.
18. Ou, S. (2023, December). Construction of Oral English Learning Model Based on Artificial Intelligence. In *International Conference on Big Data Analytics for Cyber-Physical System in Smart City* (pp. 841-854). Singapore: Springer Nature Singapore.
19. Peng, J. E., & Liang, W. (2025). Willingness to communicate with artificial intelligence (AI)? insights from tracking EFL learners' perceived acceptance and chat output. *Computer Assisted Language Learning*, 1-27.
<https://doi.org/10.1080/09588221.2025.2486147>
20. Qin, F., Sun, Q., Ye, Y., & Wang, L. (2025). A multimodal teaching model of college english based on human-computer interaction. *International Journal of Human-Computer Interaction*, 41(3), 1762-1770.
21. Rassaei, E. (2023). The interplay between corrective

- feedback timing and foreign language anxiety in L2 development. *Language Teaching Research*.
22. Roy, D. (2024). Pedagogical Restructuring of Business Communication Courses: AI-Enhanced Prompt Engineering in an EFL Teaching Context. In *Artificial Intelligence in Education: The Intersection of Technology and Pedagogy* (pp. 247-287). Cham: Springer Nature Switzerland.
 23. Savaşçı, M. (2014). Why are some students reluctant to use L2 in EFL speaking classes? An action research at tertiary level. *Procedia-social and behavioral sciences*, 116, 2682-2686.
 24. Shin, Y., Wei, S., & Vanchinkhuu, N. (2025). Digital Plagiarism in EFL Education during the AI Era: A Comparative Study of Perceptions among Learners and Instructors in Korea, Mongolia, and China. *LEARN Journal: Language Education and Acquisition Research Network*, 18(1), 594-618.
 25. Shumin, K. (2002). Factors to consider: Developing adult EFL students' speaking abilities. *Methodology in language teaching: An anthology of current practice*, 12(35), 204-211.
 26. Suban, T. S. (2021). Teaching speaking: activities to promote speaking skills in EFL classrooms. *Lectio: Journal of Language and Language Teaching*, 1(1), 41-50.
 27. Wei, W., Zhao, A., & Ma, H. (2025). Understanding How AI Chatbots Influence EFL learners' Oral English Learning Motivation and Outcomes: Evidence from Chinese learners. *IEEE Access*.
 28. Xu, Jiajin, Zhao, Chong & Sun, Mingchen. (2024). Applications of Large Language Models in Foreign Language Teaching and Research. Beijing: Foreign Language Teaching and Research Press.
 29. Xu, Jiajin, Zhao, Chong. (2024). *Roles of Large Language Models in English Language Teaching*. *Frontiers of Foreign Language Education Research*, 7(1), 3-10.
 30. Zhang, M. (2024). The Impact of Artificial Intelligence Virtual Oral Tutoring APPs on Chinese Youth's Anxiety in Oral English learning—Interview Research Based on Users of Artificial Intelligence Speaking Tutoring APPs. *Journal of Language, Culture and Education*, 1(1), 31-38.
 31. Zheng, Y. B., Zhou, Y. X., Chen, X. D., & Ye, X. D. (2025). The influence of large language models as collaborative dialogue partners on EFL English oral proficiency and foreign language anxiety. *Computer Assisted Language Learning*, 1-27.