

AI-Assisted Oral English Teaching: A Review of Generative Large Language Models

Junyao Liu*

Institute of Foreign Languages, Henan University, Kaifeng, China

**Corresponding author: Junyao Liu.*

Abstract

As generative large language models like ChatGPT are developing, their applications have become increasingly widespread and there are new technical directions for traditional English oral language teaching. However, studies have focused on writing and grammar and their systematic review of oral English teaching is still lacking. This article discusses generative large language (LLMs) models in English oral linguistic teaching, with a particular emphasis on their usage in different teaching situations and their impact on learning results. The article also compares development directions and differences between research in countries and outside the world. It is found that GenAI has seen broad applications in other ways before, during, and after class, and its users speak fluently and readily to themselves, but still suffer from language bias, content accuracy, and learning bias. Based on this, the paper suggests future development directions from the perspective of technological optimization and teaching integration in order to provide a systematic guideline for empowering English oral Language teaching with artificial intelligence.

Keywords

generative AI, large language model, oral English teaching, college classroom

1. Introduction

Nowadays, global higher education is gradually undergoing a digital transformation. Digital technology is now becoming the key step towards overcoming traditional educational resources and teaching quality improvement with intelligent teaching tools covering all teaching [1]. In this regard, generative large language models (GenAI) are rapidly advancing. Since 2022, the ChatGPT model has made its way into the education community, offering new technologies to reform and develop teaching models based on natural interaction and personalized adaptation [2]. Among other teaching situations, it is important for college oral English teaching. With the introduction of globalization, oral communication ability is of high importance for students to participate in international exchanges and increase their competitiveness at university, which is the key goal of university talent development. However, the traditional oral English course has gone through three major problems: First, large class teachers have limited teaching skills and it is hard to accommodate students of different learning levels due to a lack of personalized training resources. Second, students are afraid of making mistakes and so cannot even speak up, which reduces their effective interaction time. Third, good oral training resources are insufficient, which makes it impossible to support large-scale personalized training. These have also constrained the quality improvement of teaching. While GenAI brings many possibilities to address these

problems, there are still important gaps in research: most published studies concern GenAI in writing, grammar, and reading [1], and in oral English teaching. Moreover, few published exploratory studies in this direction do not present comprehensive reviews of GenAI application models in oral instruction, and have not discussed the integration relationship between technology and education. Based on this, this study contributes in two main ways: Theoretically, it can fill the research gaps in the field of GenAI-enabled oral teaching and refine the research landscape of this domain. Practically, it can provide systematic references for frontline teachers to optimize oral English teaching with GenAI. By systematically reviewing relevant literature, this paper clarifies the current development status of the field, laying a foundation for subsequent research and practice.

2. Literature Review

This study draws on literature retrieved from five major databases: arXiv, JSTOR, ERIC, CNKI, and Cambridge Core. A Boolean search strategy was applied using combinations of the following keywords: (“Generative AI” OR “Large Language Models”) AND (“Spoken English Teaching” OR “Oral English Teaching”). This essay included studies that adopted empirical methods (including quasi-experimental studies and case studies) or systematic reviews. Given that the technological breakthrough of generative large-language models is closely associated with the release of ChatGPT in November 2022, the search was limited to publications from November 2022 onwards; consequently, the vast majority of the literature cited herein is drawn from journal articles and conference papers published after November 2022. In the field of Artificial Intelligence-Assisted Language Learning (AIAL), I excluded studies focusing on traditional AI chatbots, early NLP tools, or those integrating multiple AI technologies, if they did not isolate the specific effects of generative large language models. This criterion was adopted to ensure the relevance and consistency of the reviewed literature.

3. Definition of Core Concepts

Generative Large Language Models (GenAI/LLMs) are deep models on a Transformer. Compared to standard AI chatbots, trained pre-trained on large multi-source corpora, these models provide deep contextual understanding, open-domain natural conversation and dynamic content generation, which can provide the user with an interactive experience [3]. In second language teaching the AI-assisted speech Dialog System (SDS) developed by this model can learn in a customized way such as simulated dialogue, role-playing and topical conversations, for different learners' languages and individual needs [4]. Oral English teaching is a teaching activity focused on developing learners' full communicative competence. In addition to using one-dimensional pronunciation or grammar training, it aims to link the three aspects of language form, meaning and communication, and jointly to increase accuracy, fluency, and total competence [5]. This is very similar to the classical Communicative Language Teaching (CLT) method that integrates grammatical, sociolinguistic and strategic competence to deal with a learner's real communicative needs [1], to achieve a complete improvement from linguistic knowledge to communication competence.

4. Research Status of Generative Large-Language Model-Assisted Oral English Teaching

At present, research on the use of generative large language models (LLMs) in oral English teaching has expanded rapidly, showing diverse application patterns and notable differences across research contexts.

4.1 Application Models

Current mainstream practices have developed applications across three instructional stages: pre-class, during-class, and after-class stages. From the functional perspective, three core application roles have been derived: 1. AI Assessor: supporting pronunciation evaluation and oral proficiency assessment; 2. AI Oral Practice Partner: facilitating interactive activities such as scenario-based dialogues and role-playing tasks; 3. AI Teaching Assistant/Learning Companion: providing students with personalized after-class exercises and instant feedback, while assisting teachers with teaching data analysis and lesson preparation [6]. There are 17 teaching activities done in 5 categories, including pre-class, in-class training, after-class teaching and training to a wide range of settings like oral conversation, role-playing, and test-based training [4]. They highlight key

benefits of LLMs (instant feedback and learning on its own) [5], provide students with their own after-school practice and instant feedback, and help teachers perform their data analysis and lesson planning [7,8].

4.2 Pedagogical Effects

On a practical level, researchers around the world have also conducted extensive experimentation: International studies have already tested the technical performance, and found that LLMs can help reduce learners' anxiety and improve their willingness to communicate. They pointed out that two main barriers when applying generative AI to language teaching. Firstly, external barriers: [5] lack of clear strategies, guidelines and ethical guidelines for generative AI applications at the level of institutionalisation, high teacher concern about data leaks and student privacy, and [9] internal barriers at the teacher level: teachers are extremely busy and rigid settings make it hard to implement these tools, and the TPACK framework (integration of technological, pedagogical and content knowledge) can tackle such internal challenges [5]. While some domestic studies started somewhat later, relevant empirical studies have continued to be done: They not only confirm that LLM can improve oral fluency and lexical complexity of non-English majors, but also implement a local human-computer collaborative teaching framework [6,8], which shows the evolution of feasibility exploration to empirical practice.

5. Issues and Limitations in Applications

Currently, generative large-language models still face several important limitations in the practical implementation of oral English teaching.

5.1 Technical Issues

At the technical level, current models still exhibit significant standard language biases: even with Refined Model 2, they default to native speaker language norms, making it difficult to adequately support the teaching needs of global English varieties. This bias has specific pedagogical consequences. First, when providing feedback, the model tends to offer standardized and abstract comments rather than personalized feedback that reflects a learner's individual voice, unique linguistic background, and cultural identity—an essential element for Global Englishes language teaching (GELT) [10]. Second, in simulated communicative interactions, the model initially may produce seemingly authentic ELF (English as a Lingua Franca) features, but as the conversation progresses, its responses gradually degrade and shift toward native-like English usage. This makes it unable to sustain linguistic diversity and authentic representation of ELF communication throughout an extended interaction [10]. Consequently, learners may receive inconsistent exposure to diverse English varieties and lack the sustained, authentic practice needed to develop communicative competence in multilingual contexts.

5.2 Pedagogical Issues

At the technical level, current models still exhibit significant standard language biases: even with Refined Model 2, they default to native speaker language norms, making it difficult to adequately support the teaching needs of global English varieties [6]. In addition, these models also raise concerns about content accuracy, as they may mislead learners [11]. Specifically, teachers need to guide students to critically evaluate AI-generated content and raise reasonable questions, preventing students from being misled by inaccurate information [6]. Moreover, these models fail to correctly identify non-verbal cues such as intonation and tone, making it hard to achieve truly immersive interactions [11]. Current large language models cannot account for the impact of non-linguistic modalities such as facial expressions and gestures on oral communication, whereas real oral communication has distinct multimodal characteristics. [12] Therefore, teachers should also help students develop the ability to recognize and choose multiple modalities, especially coordinating body language with verbal language for more effective and expressive oral English communication [6]. At the pedagogical level, learners tend to over-rely on the models, leading to cognitive inertia and weakening their thinking abilities, which in turn diminishes their innovative and autonomous thinking skills [11]. To prevent students from falling into shallow thinking due to over-reliance on AI, teachers must guide them to critically evaluate model outputs and encourage independent thinking [6].

6. Future Directions

To address the aforementioned problems, future research and practice can be extended from the following directions: Technically, educational models and corpora can be developed via API tuning, and virtual reality (VR) can be employed to build immersive interactive environments that enable learners to conduct self-assessment of their oral proficiency [6,11]. Pedagogically, future research can explore how to cultivate teachers' natural prompting abilities, design personalized courses for cross-cultural communication, and construct a human-computer collaborative teaching model-all aimed at making students' learning more relevant to their needs and contexts [6].

7. Conclusion

In this work, we review current research on GenAI for oral English teaching in Chinese universities. Given teaching resources and pedagogical models, traditional oral English Teaching in Chinese Universities has long failed to satisfy personalized learning needs. In recent years, GenAI has provided new routes to this task and a full-process application system including pre-class, during-class and after-class has been developed. In terms of research progress, international work has focused on technical efficiency and application issues of generative large language models, and domestic studies have focused on local validation and local teaching model building. However, the technology still faces standard language bias, poor content accuracy, and potential overload of students. It needs to optimize the technology, incorporate VR tools, and build a human-computer collaborative teaching for quality and efficiency.

References

- [1] Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, 1(1), 1–47. <https://doi.org/10.1093/applin/1.1.1>
- [2] Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., ... Amodei, D. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems*, 33, 1877–1901. <https://arxiv.org/abs/2005.14165>
- [3] Chen, B. (2025). The logic, dilemma and practical path of AI enabling personalized and accurate teaching of spoken English. *China Journal of Multimedia & Network Teaching (Mid-month Issue)*, 87–90.
- [4] Ge, N., & Zheng, C. P. (2025). A systematic review of artificial intelligence chatbots supporting oral English teaching. *Foreign Language Education*, 46(3), 26–34. <https://doi.org/10.16362/j.cnki.cn61-1023/h.2025.03.013>
- [5] Lee, S., Choe, H., Zou, D., & Jeon, J. (2026). Generative AI (GenAI) in the language classroom: A systematic review. *Interactive Learning Environments*, 34(1), 335–359. <https://doi.org/10.1080/10494820.2025.2498537>
- [6] Wu, L. J. (2025). An empirical study of large language models enabling oral English teaching. *Modern Educational Technology*, 35(10), 33–41. <https://doi.org/10.3969/j.issn.1009-8097.2025.10.004>
- [7] Su, Q. (2024). Efficacy assessment of large language models application in L2 teaching. *Foreign Language World*, (3), 35–42.
- [8] Wang, C. X., Peng, L. H., & Jiang, X. (2026). An action research of the effectiveness of English teaching with AI support in universities. *Foreign Languages and Literature*, (1), 210–220. <https://doi.org/10.26922/j.cnki.waiguoyuwen.2026.01.017>
- [9] Shi, X. (2023). Advantages, challenges and prospects of ChatGPT in oral English teaching. *Transactions on Social Science, Education and Humanities Research*, 4, 123–127.
- [10] Lee, S., Jeon, J., McKinley, J., & Rose, H. (2024). Generative AI and English language teaching: A global Englishes perspective. *ELT Journal*, 79(1), 12–24. <https://doi.org/10.1093/elt/ccae052>

- [11] Zhang, Y. (2025). Application of large-language-model-based intelligent study assistant in college spoken English learning: A case study of T college. *International Journal of Language and Education Research*, 7(1), 116-139.
- [12] Wu, L. J. (2013). An empirical study on the impact of multimodal English teaching on college students' multiliteracy ability. *Modern Educational Technology*, 23(10), 82-86. <https://doi.org/10.3969/j.issn.1009-8097.2013.10.016>

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Open Access

This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

