

The Effects of AI Oral Tools on Oral Proficiency Among English Teacher Candidates: A Quasi-Experimental Study

Yi Wang^{1, a}

¹English Education, Hangzhou Normal University, Hangzhou, Zhejiang 311121, China

^a1684519870@qq.com

Abstract

Research on the Effect of AI-Assisted Speaking Practice on the General Level of English-Speaking Ability for Teacher Trainees: A Quasi-Experimental Study. Over the next 12 weeks, this experiment will compare the learning results of students in traditional classes with those of students who have participated in the combined model of daily AI-assisted exercises. Some kinds of blind assessment and statistics will be used. According to the above data, AI oral tools can improve students' speaking fluency, phonetics and grammar accuracy, vocabulary and application in actual teaching situations. Problems in the old method of oral teaching, such as limited practice opportunities, slow feedback and an absent professional communication environment, can also be resolved by the tools mentioned above. The results offer some practical suggestions for self-study and practice by English-teaching candidates and can also help to strengthen the curriculum of university oral-English courses.

Keywords

AI oral tools; English teacher candidates; Oral proficiency; Quasi-experimental study; Teaching interaction.

1. Introduction

With the continuous improvement in the quality of basic English education, the oral proficiency of English teacher candidates has become a crucial factor that directly affects the quality of teaching and students' language learning. With the development of artificial intelligence, many intelligent oral devices have now appeared that can solve some of the problems of traditional oral teaching. Based on the actual situation, custom training routes and scenario-based simulations have been added to solve many problems in teacher education.

Many English teacher candidates still lack speaking opportunities, receive late or irregular feedback, and do not have access to practical teaching communication environments. There are now deficiencies in the professional oral skills needed for teaching in the classroom. Although AI has been widely applied in general language education, there are relatively few studies focusing on English teacher candidates. Most of the attention is on ordinary college students, and there is little integration of profession-specific tasks such as classroom teaching, teacher-student interaction and instructional discourse. Most research designs are short-term and do not have the advantages of long-term observation, repeated evaluation and blinding. Therefore, very few practical and scalable training frameworks have been developed for this group in particular.

Given the above deficiencies in the current research, this paper will investigate the actual effects of AI oral tools on the oral performance of English teacher candidates and put forward some practical implementation paths. The three main research problems are:

Can AI oral tools help to improve the general oral ability of English teacher candidates?

How good are the above tools in improving the quality of teaching-oriented interactive activities?

What should a Usable AI-assisted Oral Training Framework for English Teacher Candidates look like?

The first is to expand the scope of AI-assisted teacher education and to enhance the application of communicative language ability and technology acceptance models in professional preparation. In practice, it can provide a reasonable self-study plan for teacher candidates, reduce the feedback burden on university instructors, help improve the oral English curriculum, and thus support the supply of more capable English teachers for basic education.

2. Literature Review

2.1. International Research

Most of the foreign research on AI-assisted oral instruction is still in its early stages and has not accumulated much evidence. Tao and others (2025) [1] have found that generative AI and VR can help students learn self-regulation and reflection. Zhang (2025) [2] carried out a 16-week longitudinal case study and found that AI oral applications can offer personalised and timely correction of pronunciation and grammar. Ye and others (2025) [3] have found that student' trust in AI chatbots is positively related to their wish to speak. Wan and Moorhouse (2025) [4] have systematically arranged AI conversational partners and their functions for oral practice. Given the communication language ability, technology acceptance theory and self-regulated learning, most foreign studies are based on experiments. However, very few have focused specifically on pre-service teachers and have not been tailored to particular situations.

2.2. Domestic Research

At home, people have been studying the effects, teaching models and student motivation of AI oral tools. Jianhao Wu and others (2024) [5] found that generative AI can improve children's speaking fluency and grammatical structure. Lingjuan Wu (2025) [6] put forward a model of human-computer cooperation in oral instruction. Nan Ge and Chunping Zheng (2025) [7] pointed out that intelligent chatbots can help people practise speaking in private.

Fang Huang and Zhaojie Tang (2026) and Bin Zou and Chenghao Wang (2026) [8] [9] found that the three factors of perceived enjoyment, satisfaction and usefulness can motivate people to keep using a product. Yinxing Jin and Fengjuan Sun (2025) [10] found that AI can increase happiness and self-confidence and reduce anxiety about speaking. Yangyang Li and Jing Wang (2023 [11]), Jian'er Zhong (2020 [12]) and others have been carrying out research on virtual simulation and curriculum reform for English teacher candidates. No empirical studies have been carried out on AI oral tools and training scenarios for some professions yet.

One of the main differences between the two is that foreign studies generally focus on learning mechanisms and long-term changes in behaviour, whereas domestic research has paid more attention to short-term gains in fluency, accuracy and learning efficiency. Few studies have connected the improvement of language skills with professional growth in a single design, and this is what the current study aims to do.

2.3. Identified Gaps

Firstly, most of the previous studies have used general undergraduate students as the subject group and have not focused on English teacher candidates. Second, most AI applications are still general language exercises and have paid little attention to teaching interaction, classroom management talk and instructional delivery. Thirdly, there is no uniformity in the empirical design; most are short-term projects, and studies that have carried out longitudinal tracking,

multi-wave assessments and blind teacher ratings are still few. Fourth, there are no actual AI training plans for teacher candidates.

2.4. Innovations of the Present Study

The four additions to the previous research are:

1. Population Focus: English teacher candidates are the targets, and the teaching standards require it.
2. Content Design: Basic language skills training and teaching scenario simulation.
3. Methodological Rigour: In the past 12 weeks, pre-tests, two interim observations, post-tests and blind ratings by experienced teachers have all taken place.
4. Practical Output: It is a ready-to-use AI oral training plan with good academic credentials and classroom suitability.

3. Research Methods

3.1. Participants and Recruitment

Thirty sophomore students in the English Education major of a local teacher training college were selected from the two regular sophomore classes. All had finished the basic oral English course and showed similar initial oral proficiency. All the subjects were eligible for the pre-test. One class was set up as the experimental group, and the other as the control group; both had 15 students. Recruitment was voluntary. In advance of conducting the interview, participants were informed of the reasons for the interview, any need to record, etc.

3.2. Research Design

A Quasi-Experimental Design was employed. The teaching content, contact time and instructors in the experimental and control groups were kept the same to rule out other reasons. It was 12 weeks long. Both groups had regular in-class oral instruction and 30 minutes of daily AI-assisted oral practice. Only the control group received regular oral instruction and did not have structured AI practice. Four oral proficiency tests were conducted at the beginning (pre-test), week 4, week 8 and week 12 of the program (post-test) to track progress over time.

This quasi-experimental design was chosen because it allows for comparison under natural teaching conditions without excessive disruption to normal class arrangements. Given the use of intact classes, cluster-level statistical analysis was employed to account for nested classroom effects and ensure the validity of the results. A strict randomized controlled trial would be difficult to implement in this educational setting due to scheduling and administrative constraints. Thus, a quasi-experimental design provides a reasonable balance between internal validity and practical feasibility.

3.3. Variable Definition

3.3.1. Independent variable

AI-supported oral practice, including foundational skill drills and teaching-scenario simulation.

3.3.2. Dependent variable

Overall oral proficiency, covering pronunciation and intonation, fluency, vocabulary and grammar, discourse coherence, interactional strategies, and teaching-scenario appropriateness.

3.3.3. Control variables

Instructor, teaching content, class schedule, assessment rubric, testing timing, and initial oral proficiency.

3.4. Research Instruments

AI oral platforms: Duolingo, ChatGPT, and Doubao were selected for their reliable speech recognition, real-time feedback, and scenario-based dialogue functions.

Oral rating rubric: Based on the *Chinese English Proficiency Scale* and Bachman & Palmer's communicative language ability model, a two-dimensional scoring rubric was developed:

Basic Linguistic Dimensions (60pts)	
Pronunciation and Fluency (20)	
Lexical Resource and Accuracy (20)	
Grammar and Logic (20)	
Professional Interactive Dimensions (40pts)	
Discourse Coherence (15)	
Interactive Strategies (15)	
Teaching Scene Adaptability (10)	

Blind Rating Procedure: Three experienced English teachers with over five years of teaching experience independently rated all the recordings without knowing the group. The final analysis is the mean score.

Data Management: An audio recorder, Excel spreadsheets and SPSS 26.0 were used to collect, store and analyse the data.

3.5. Procedure and Implementation

1. Preparation (Weeks 1-2): Finalise the research plan, prepare assessment materials, train raters, recruit subjects and complete the group assignment of intact classes.
2. Pre-test (Week 3): Conduct the first round of oral examinations and confirm group homogeneity.
3. Intervention (Weeks 4-12): The control group received the regular oral English curriculum. Both groups had the same curriculum, but the experimental group added 30 minutes of daily AI training to improve pronunciation correction, fluency, vocabulary and grammar, classroom instruction scripting and teacher-student interaction simulation.
4. Interim Assessments (Weeks 4 and 8): Check the progress, record the scores and audio, and adjust the practice tasks accordingly.
5. Post-test (Week 12): Conduct the final oral examination and collect all data.
6. Data Processing: Organise the rating data and analyse it.

3.6. Data Analysis

SPSS 26.0 was used for the statistics above. Calculate and present the mean and standard deviation of the data at each time point. Independent-samples t-tests and paired-sample t-tests were used to compare the differences in the experimental group and the control group. Cluster analysis was used to correct for the intra-class correlation caused by the grouping of intact classes and thus ensure the reliability of the statistical results. Statistical significance was set as $p < 0.05$.

4. Expected Results and Analysis

4.1. Expected Group Differences

Based on the pre-test results, the total oral scores of the experimental group and the control group are not significantly different; thus, these two groups are considered equivalent at the

start. By the end of week 4 and week 8, the experimental group will show better progress in fluency and pronunciation accuracy. Based on the results of the post-test, it is expected that the experimental group will score higher than the control group in many areas, such as pronunciation and intonation, lexical diversity, grammatical accuracy, and appropriateness of teaching-related interactions. Although there will be some improvement in the control group after 12 weeks, the degree of increase will be relatively small, especially for interactive and professionally specific performances.

4.2. Expected Variable Relationships

The frequency of AI application will help improve the students' oral language skills; thus, continuous and timely feedback will be provided in all sorts of circumstances. It is hoped that the improvement in teaching-scenario interaction will exceed the increase in basic language skills, and therefore AI will be more beneficial for professional oral communication. Self-efficacy and the regularity of practice are also expected to predict the final score; that is, more willing students will perform better.

4.3. Hypotheses to Be Tested

1. AI-supported practice significantly improves overall oral proficiency among English teacher candidates.
2. AI produces a significantly larger effect on teaching-scenario interactive performance than traditional instruction alone.
3. Continuous long-term AI practice leads to more stable learning gains than short-term intervention.

4.4. Plausible Explanations

AI creates a low-anxiety environment that encourages frequent speaking and reduces performance pressure. Real-time phonetic and grammatical feedback helps correct errors immediately and systematically. Personalized practice pathways accommodate diverse learner levels. Teaching-scenario simulation strengthens targeted professional discourse. In contrast, traditional instruction is constrained by limited speaking turns, delayed feedback, and fewer contextualized tasks, resulting in more modest improvement.

5. Conclusion

This 12-week quasi-experimental study examines how AI oral tools shape the oral proficiency of English teacher candidates. Therefore, based on the above, a pre-test-midterm-post-test design will be used to conduct blind instructor ratings and statistical analysis of the effect of structured AI practice. Based on the above, it is expected that AI tools can help students improve the fluency, accuracy, range of vocabulary and grammar, and interaction ability in teaching situations.

With the help of artificial intelligence, various problems in teaching English at the primary school level have been addressed, and the results of this study will provide a reference for other teachers. It offers a good self-training model for teacher candidates and provides information on the reform of curriculum in university oral English classes to support high-quality teacher education.

In addition to the first point, another is that AI tools help teacher candidates establish a professional way of speaking in their daily work, not just enhance their language skills. Therefore, the aim is to cultivate students with professional literacy in line with the requirements of intelligent education at present.

Some shortcomings have been omitted. The sample is small and taken from only one university; therefore, it may not be generalisable. Although it has been 12 weeks, it have not yet seen any

long-term effects or sustainability. Compliance with AI practice is the result of students' self-discipline and will vary slightly. In the future, this experiment will expand the number of institutional samples, carry out multi-site comparisons, lengthen the follow-up period, collect all kinds of data, and thus improve both the framework and the external validity.

References

- [1] Lei Tao, Yanjie Song & Jiachen Fu. (2026). Exploring students' self-regulated learning behavioural patterns and perceptions in an English speaking task within a generative AI-supported immersive VR environment. *Computers & Education*, 242, 105515-105515. <https://doi.org/10.1016/j.COMPEDU.2025.105515>.
- [2] Xiudi Zhang. (2025). Enhancing oral English learning through AI: a case study on the impact of AI-driven speaking applications among Chinese university students. *Frontiers in Psychology*, 16, 1595818-1595818. <https://doi.org/10.3389/FPSYG.2025.1595818>.
- [3] Long Ye, Jie Du & Zhaohui He. (2025). Willingness to communicate with AI chatbots in English: the role of personality and trust among Chinese college students. *Disability and Rehabilitation: Assistive Technology*, 1-16. <https://doi.org/10.1080/17483107.2025.2572532>.
- [4] Yuwei Wan & Benjamin Luke Moorhouse. (2025). Using Call Annie as a Generative Artificial Intelligence Speaking Partner for Language Learners. *REL C Journal*, 56(2), 489-498. <https://doi.org/10.1177/00336882231224813>.
- [5] Jianhao Wu, Wanting Zhou & Chao Cao. (2024). An empirical study on generative artificial intelligence-enabled oral English teaching. *China Educational Technology*, 4, 105-111.
- [6] Linjuan Wu. (2025). An empirical study on large language models empowering English speaking teaching. *Modern Educational Technology*, 35(10), 33-41.
- [7] Nan Ge & Chunping Zheng. (2025). A systematic review of studies on intelligent chatbot-supported English speaking teaching. *Foreign Language Learning*, 46(3), 56-63. <https://doi.org/10.16362/j.cnki.cn61-1023/h.2025.03.013>
- [8] Fang Huang & Zhaoqi Tang. (2026). AI-enabled oral English learning: A study on perceived enjoyment and continuance intention. *Journal of Xi'an International Studies University*, 34(1), 53-59. <https://doi.org/10.13632/j.cnki.cn61-1457/h.2026.01.009>
- [9] Bin Zou & Chenghao Wang. (2026). Generative artificial intelligence as a pathway for personalized English speaking learning among college students. *Foreign Language Learning*, 1, 61-69. <https://doi.org/10.16263/j.cnki.cn23-1071/h.2026.01.010>
- [10] Yinxing Jin & Fengjuan Sun. (2025). A study on generative AI-based emotional intervention in college students' English speaking anxiety. *Foreign Language Education Research Frontiers*, 8(3), 97-109. <https://doi.org/10.20083/j.cnki.fleic.2025.0031>
- [11] Yangyang Li & Jing Wang. (2023). A pathway to improve pre-service English teachers' oral expression ability in the context of virtual simulation technology. *English Square*, 20, 86-91. <https://doi.org/10.16723/j.cnki.yygc.2023.20.029>
- [12] Jian'er Zhong. (2020). Effective approaches to cultivating normal university students' oral English proficiency in the context of the national teacher qualification examination. *Journal of Jiamusi Vocational College*, 36(8), 117-118.