Assessing the Efficacy of AI-Powered Speaking Apps in China's Mainland Middle Schools: A Mixed-Methods Study on Oral Accuracy and Fluency

Shanyuan Zhu[1, a]

[1]The University of Nottingham, Nottingham, England

[a]15118888004@163.com

Abstract

English plays an increasingly crucial role in cross - cultural communication and education, especially for secondary students in China's mainland, where oral proficiency has drawn significant attention due to high - stakes examinations like zhongkao. Conventional English language teaching methods face challenges such as public speaking anxiety, delayed feedback, and limited practice opportunities. AI - powered speaking tools emerge as a potential solution, integrating real - time feedback and adaptive learning technologies. For example, students can receive feedback or obtain suggestions instantly following their completion of speaking exercises. This mixed - methods study aims to evaluate the effectiveness of these tools in enhancing oral accuracy, such as pronunciation, and fluency among middle school students. A cohort of 32 ninth - grade students preparing for zhongkao participated in a structured questionnaire survey. The results show that 56% of the students considered AI tools beneficial for pronunciation refinement and fluency enhancement, although improvements in grammatical precision and lexical diversity were less evident. The main advantages highlighted by participants include technological affordances like instantaneous feedback from speech recognition algorithms and pedagogical flexibility like self - paced practice outside class. However, barriers such as technical limitations (device compatibility and network instability) and economic constraints (high subscription costs) also exist, which may impede sustained engagement.

Keywords

AI-powered tools, fluency, accuracy, oral English.

1. Introduction

English, as a global lingua franca (Elizabeth, 2022), is used in cross-cultural communication more and more (Wang, 2024). As a result, speaking English is becoming increasingly important in education, especially for secondary students in grades 7–12 in China's mainland. Oral proficiency has garnered substantial societal attention, driven predominantly by the exam-oriented education system. High-stakes assessments, such as the senior high school entrance examinations (zhongkao), mandate oral English components, pressuring students to prioritize spoken fluency and accuracy. However, conventional pedagogical approaches to English language instruction face criticism (Yang, 2020) for perpetuating challenges like public speaking anxiety (Men, 2024), delayed corrective feedback, and limited practice opportunities (Elizabeth, 2022). Moreover, traditional English teaching tends to place excessive emphasis on grammar and vocabulary, overlooking the significance of oral practice.

Emerging AI-powered speaking applications offer a potential remedy. By integrating real-time feedback and adaptive learning technologies, these tools may address systemic gaps in oral skill development. This study adopts a mixed-methods design to evaluate the effectiveness of AI-

driven apps in enhancing oral accuracy like pronunciation and fluency such as speech rate, coherence among middle school students.

A cohort of 32 ninth-grade students preparing for zhongkao participated in a structured questionnaire survey. Findings revealed that 56% perceived AI tools as beneficial for pronunciation refinement and fluency enhancement, while improvements in grammatical precision and lexical diversity were less frequently reflected. Participants highlighted two primary advantages: (1) technological affordances, such as instantaneous feedback enabled by speech recognition algorithms, and (2) pedagogical flexibility, allowing self-paced practice outside classroom hours. Conversely, barriers included technical limitations such as device compatibility issues and network instability, and economic constraints like high subscription costs, which deterred sustained engagement.

2. Literature Review

In the context of educational informatization, significant emphasis ought to be laid on the reform of English curriculum teaching and learning. The Ministry of Education has recognized the limitations of traditional teaching methods and aims to integrate AI technology with exam-oriented education to enhance students' English-speaking abilities. AI that incorporate sophisticated Natural Language Processing (NLP) architectures has found numerous applications in school-based foreign language education. The pedagogical transformation of English language education can be substantially enhanced through the integration of AI-driven machine (Fitria, 2021).

The integration of AI-driven technologies presents multiple advantageous implications in academic research contexts. First of all, AI-powered tools help to improve phonetic accuracy (López-Minotta et al., 2025), phonological accuracy and conversational fluency (Chen et al., 2025; Du et al., 2024). In addition, AI-driven educational platforms offer real-time feedback mechanisms that enable continuous assessment and inform instructional adjustments for both educators and learners (Zhang et al., 2023). Moreover, some students are reluctant to speak English voluntarily because they fear making mistakes or getting unfavorable feedback from peers or teachers. The integration of AI-driven platforms into spoken English practice demonstrates stress-alleviating effects (Du et al., 2024) and foster incremental confidence-building through sustained engagement (Crompton et al., 2024). It significantly boosts student motivation and involvement (López-Minotta et al, 2025).

While the incorporation of these tools holds great promise, recent studies have shed light on a series of challenges that cannot be overlooked. One significant area of concern is the prevalence of technical glitches. As Du et al. (2024) have noted, problems such as recognition failures, inaccurate responses, and internet connectivity issues can have a detrimental impact on the user experience. Another limitation lies in the personalization capabilities of current AI systems. Chen et al. (2025) point out that restricted training datasets and resource scarcity impede the provision of personalized learning materials. In today's dynamic educational landscape, students require up-to-date and relevant resources to engage in in-depth discussions and stay informed. Finally, Crompton et al. (2024) highlight the emotional void in AI interactions. Learning is not just an intellectual process but also an emotional one. Human interaction often provides emotional support, encouragement, and a sense of connection, which are lacking in AI-driven exchanges. This emotional deficiency may affect students' motivation and engagement, as well as their overall well-being during the learning process.

In conclusion, while AI-powered tools offer potential benefits, these limitations call for further research and development to overcome the challenges and ensure that these tools can be effectively integrated into the academic sphere. Most previous research mainly focused on adult learning or college students, whose learning motivation stems from personal interests, rather

than junior high school students, who are motivated by exams. The most prominent benefit of this application lies in the combination of AI technology with exam - oriented learning and the attention given to learners' own feelings and feedback after using these AI - powered tools.

3. Method

This study adopted a mixed-method approach to evaluate the effectiveness of AI-driven speaking applications in improving oral accuracy and fluency. To learn more about the efficacy and perceptions of AI-powered speaking aids, questionnaires were disseminated. Descriptive statistics were used to evaluate quantitative data, while thematic coding was used to categorize qualitative replies. Students in the ninth grade of Guangzhou, who reside in an educational environment characterized by widespread oral English application and pervasive AI-driven technologies, participated in a survey study. Beginning in the ninth grade, these students are mandated to utilize Gaofenshuo, an AI-powered language platform, to complete weekly homework assignments designed to facilitate regular oral English practice. It integrates an intelligent speech evaluation system to facilitate extensive input-output language practice, thereby promoting language internalization. As a result, most of them has used AI-powered tools before and is familiar with both conventional and AI teaching methods. The survey was composed of three distinct components:

A. Demographic Data: Basic information such as language proficiency level, and app usage frequency was collected. This data provides a foundational understanding of the characteristics of the participants, which can potentially influence their experiences and evaluations of the AI-driven speaking applications.

B. Core Metrics: a) Targeted Efficacy: This metric aimed to measure the degree to which the app meets specific learning needs, such as pronunciation refinement and grammar correction. By asking "To what extent does the app address specific learning needs (e.g., pronunciation refinement, grammar correction)?", we sought to understand the app's performance in fulfilling the learners' individualized requirements. b) Proficiency Impact: This dimension investigated whether there had been a noticeable improvement in oral accuracy (including pronunciation and grammar) after the adoption of the app. The question "Has your oral accuracy (pronunciation/grammar) demonstrably improved post - app adoption?" was designed to directly assess the impact of the app on learners' language proficiency.

C. Open - Ended Analysis

Two exploratory questions were included to elicit nuanced and detailed perspectives. a) Comparative Advantages: By asking "What do you perceive as the most significant benefit of AI - driven platforms over traditional language learning methods?", we aimed to uncover the unique selling points of AI - based applications from the learners' viewpoints. This could provide valuable insights for further development and promotion of such tools. b) Limitation Identification: The question "What technical or pedagogical constraints have you encountered during AI app usage?" was intended to identify the pain points and areas for improvement in the current AI - driven speaking applications.

The quantitative dataset was subjected to descriptive statistical analysis via Microsoft Excel, calculating central tendency measures (mean), dispersion metrics (standard deviation), and distribution characteristics (frequency distributions). Qualitative responses from open-ended questions underwent thematic analysis utilizing a dual-phase coding strategy: initial deductive coding based on established theoretical frameworks from prior educational research, followed by inductive coding to uncover novel themes emerging from participant narratives regarding learning experiences and engagement dynamics.

4. Statistical Analysis

Through quantitative and qualitative data, this study examined the usage patterns of AI-powered speaking applications among Chinese middle school students, their perceptions of increases in spoken accuracy and fluency, and their opinions on the benefits and limitations of these tools. A full set of usage metrics was covered by the 32 valid questionnaires that were gathered. Despite its small size, the sample is representative since it includes students who have been using AI-powered tools to improve their speaking for at least six months, indicating high persistence and stability in user groups. Therefore, it gives us a variety of viewpoints on AI-powered tools and facilitates the discussion of whether or not these tools may help people become more accurate and fluent speakers.

The primary purpose of use is exam preparation (81.25%), indicating a student-dominated user group with clear learning goals. Completing teacher-assigned tasks (40.63%) also shows the tool's role in classroom support.

Table 1. Improvement in Oral Accuracy (5-point scale)

Option	Subtotal	Percentage
Significantly improved (5)	5	15.63%
Somewhat improved (4)	13	40.63%
No change (3)	14	43.75%
Slightly declined (2)	0	0.00%
Significantly declined (1)	0	0.00%
Mean	3.72	
Standard Deviation	0.89	

As demonstrated by the data presented in Table 1 , approximately 56.26% of respondents observed enhanced accuracy levels (4-5points), whereas 43.75% reported no discernible improvement, highlighting the existence of significant individual variability in tool efficacy. Empirical evidence derived from perspective-based analysis and interviews with students who perceived no significant impact reveals that the factors influencing their decisions can be categorized as follows. On the one hand, varied perceptions regarding the different understanding of AI tool effects could contribute to divergent responses in survey data. While some students perceive a one-point increase as representing progress, a segment of the student body considers such incremental change inconsequential in performance evaluation. On the other hand, motivation may be the cause of individual differences. For instance, some students may make significant development if they use speaking tools to complete oral practice thoroughly and give themselves further oral tasks, while others who did not take the task seriously may see just modest progress. Lastly, students' divergent English proficiency levels may shape their preferences, as AI-powered tools demonstrate greater efficacy in assisting students with limited oral proficiency compared to those already possessing advanced speaking competencies.

Table 2 highlights the usefulness of AI in speech training and fluid expression by revealing that the most frequently reported gains are pronunciation (75%) and fluency (71.88%). Engwall and Balter (2007) claim that an AI-powered tool with real-time automatic feedback can fix pronunciation in English. Similar to this, several students also stated that they can improve their accuracy, especially in pronunciation by receiving quick feedback that identifies their mistakes. However, the poll reveals that AI-powered products have less of an impact on vocabulary and grammar. It's possible that the primary goal of AI-powered tools is to increase

pronunciation accuracy. The AI-driven tools provide less grammar comments and no additional words to help students grow their word banks.

Table 2. Most Helped Aspects

Option	Subtotal	Percentage
Pronunciation	24	75.00%
Fluency	23	71.88%
Intonation/Pace	16	50.00%
Grammar	11	34.38%
Vocabulary	7	21.88%
Others	2	6.25%

In response to the question of the open-ended poll, "What do you consider to be the most significant advantage of AI speech applications over traditional pedagogical approaches?" Technology and the learning component are the two primary benefits, according the respondents. First of all, AI-driven tools are convenient and flexible due to the rapid feedback provided by the latest technology. There are 37.5% of students, 12 responses, praised the speaking apps for enabling seamless daily practice. Students emphasized it is convenient for daily use and some of them stress AI-powered tools' unrestricted availability such as being used anytime and anywhere. Secondly, Technical features such as "Automatic scoring function" and "Accurate pronunciation" were highlighted by 15.6% students, reflecting advancements in speech recognition and real-time correction. Pokrivcakova (2009) also pointed out a similar finding, stating that AI-driven tools can offer learners reliable feedback within just a few seconds. This immediate feedback functionality is beneficial for learners as it promotes their progress. Moreover, according to 12.5% of students, the apps support independent problem-solving ("Independent problem-solving") and self-directed learning ("No restrictions, learn freely"), which is in line with individualized educational needs.

In analyzing the technical and pedagogical limitations reported in the survey "What technical or pedagogical constraints have you encountered during AI app usage?" The research highlighted obstacles across two main categories, which are technical constraints and cost barriers. On the one hand, 56.25% students mentioned that they met technical constraints such as device or network Issues. Frequent complaints included "Device lag" and "Audio interference," pointing to hardware and software compatibility and connectivity problems. Students also critiqued "Inaccurate recognition" and "Lack of personalized plans" from the standpoint of functional limits, pointing out areas where precision and customisation were lacking. The company ought to think about raising the degree of technology.

However, another major obstacle to students using AI-powered tools is cost. High subscription rates and premium service costs, such as pricey VIP, were condemned by 34.38% of students as being barriers to continued use. It implies that the government and businesses can work together to lower prices or offer some free areas for students to utilize in order to address their financial concerns.

5. Conclusion

In conclusion, the survey reveals that AI-driven speaking applications offer two notable advantage —including measurable improvements in students' oral proficiency (pronunciation accuracy and speaking fluency)—but also present two significant drawbacks related to technical limitations and cost barriers. Technical constraints emerged as critical barriers, with

recurring issues of delayed system responses and speech recognition inaccuracies affecting user experience. Financial accessibility remains another challenge, as subscription costs limit equitable access despite the demonstrated educational value of these tools. One third of students denounced premium service fees and hefty subscription rates as deterrents to ongoing use. It suggests that not only can the companies put in place freemium models and specially design adaptive courses while keeping core functions affordable, but the government can also provide funding for company to reduce operating costs and provide some free sections for students to use in order to alleviate their financial worries.

There are some limitations in this paper. Firstly, the study's scope is constrained by its focus on exam-preparation contexts and a relatively small sample (n=32), which may limit its applicability to non-academic or adult language learners. Future research should prioritize longitudinal studies with larger, demographically diverse cohorts to explore AI speaking apps' role in fostering lifelong language learning. Additionally, future researches can adopt cross-cultural comparative frameworks. Conduct parallel experiments in diverse sociolinguistic contexts to closely examine how cultural variables interact with the efficacy of AI speaking apps. Moreover, comparative methodology can be used in future research. Establish controlled experimental designs that contrast AI-based interventions with traditional pedagogical approaches. This will help clarify their relative effectiveness in dimensions such as user engagement, fluency development, and metalinguistic awareness. By integrating these strategies, a more in-depth and comprehensive understanding of AI speaking apps can be achieved.

References

[1] Chen, Z., Song, Y., Xu, L., Liu, X., & Zu, G. (2025). The impact of artificial intelligence-assisted teaching on the improvement of oral ability of English majors. Modern Linguistics, (1), 260–268. https://doi.org/10.12677/ml.2025.131038

[2] Crompton, H., Edmett, A., Ichaporia, N., & Burke, D. (2024). AI and English language teaching: Affordances and challenges. British Journal of Educational Technology. https://doi.org/10.1111/bjet.13460

[3] Du, J., & Daniel, B. K. (2024). Transforming language education: A systematic review of AI-powered chatbots for English as a foreign language speaking practice. Computers and Education: Artificial Intelligence, 6, 100230. https://doi.org/10.1016/J.CAEAI.2024.100230

[4] Elizabeth, F., Quito, P., Magdalena, I., Ullauri, M., & Riobamba, E. (2022). UNIVERSIDAD NACIONAL DE CHIMBORAZO FACULTAD DE CIENCIAS DE LA EDUCACIÓN, HUMANAS Y TECNOLOGÍAS CARRERA DE PEDAGOGÍA DE LOS IDIOMAS NACIONALES Y EXTRANJEROS Title of the project "Self-confidence and its relationship with the development of English oral communication competence" Thesis Tutor.

[5] Engwall, O., & Bälter, O. (2007). Pronunciation feedback from real and virtual language teachers. Computer Assisted Language Learning, 20(3), 235–262. https://doi.org/10.1080/09588220701489507

[6] Fitria, T. N. (2021). Article in ELT Echo The Journal of English Language Teaching in Foreign Language Context: https://doi.org/10.24235/eltecho.v6i2.9299

[7] Liu, S.-C., & Hung, P.-Y. (2016). Teaching Pronunciation with Computer Assisted Pronunciation Instruction in a Technological University. Universal Journal of Educational Research, 4(9), 1939–1943. https://doi.org/10.13189/ujer.2016.040902

[8] López-Minotta, K. L., Chiappe, A., & Mella-Norambuena, J. (2025). Implementation of Artificial Intelligence to Improve English Oral Expression. Multidisciplinary Journal of Educational Research, 15(1), 43–71. https://doi.org/10.17583/remie.16188

[9] Men, Q. (2024). Exploration of the Application of Artificial Intelligence in High School English Speaking Teaching. Communications in Humanities Research, 35(1), 133–137. https://doi.org/10.54254/2753-7064/35/20240035

[10] Ministry of Education of the People's Republic of China. (2022). English Curriculum Standards for Compulsory Education (2022 Edition). Beijing Normal University Press.

[11] Pokrivcakova, S. (2019). Preparing teachers for the application of AI-powered technologies in foreign language education. Journal of Language and Cultural Education, 7(3), 135–153. https://doi.org/10.2478/jolace-2019-0025

[12] Wang, J., Zhang, J., Chen, L., & Dai, J. (2024). Scientific Journal of Intelligent Systems Research Study of Artificial Intelligence-Assisted English Oral Teaching.

[13] Yang, J. (2020). Using Contextualized Materials to Teach English Grammar. https://repository.usfca.edu/capstone

[14] Zhang, X., Sun, J., & Deng, Y. (2023). Design and Application of Intelligent Classroom for English Language and Literature Based on Artificial Intelligence Technology. Applied Artificial Intelligence, 37(1), 2216051. https://doi.org/10.1080/08839514.2023.2216051