

EXPLORING PATHWAYS OF UNIVERSITY STUDENTS' PERSONALISED ENGLISH-SPEAKING LEARNING WITH GENERATIVE ARTIFICIAL INTELLIGENCE

LAN, Y.^{1,2} – ZAMRI, N.^{3*}

¹ *Akademi Pengajian Bahasa, Universiti Teknologi MARA, Selangor, Malaysia.*

² *Anyang Normal University, Henan, China.*

³ *Akademi Pengajian Bahasa, Universiti Teknologi MARA Kampus Dengkil, Selangor, Malaysia.*

**Corresponding author
e-mail: norazrin5893[at]uitm.edu.my*

(Received 05th March 2026; revised 20th June 2026; accepted 08th July 2026)

Abstract. English-speaking instruction has long been plagued by insufficient input, lack of interaction, delayed feedback, and emotional anxiety, leading to unsatisfactory outcomes. Generative Artificial Intelligence (GenAI) offers a new solution to these challenges by enabling real-time, adaptive conversational practice. This study systematically examined the potential effectiveness of GenAI in enhancing learners' English-speaking proficiency using a mixed-methods approach that combined pre- and post-test speaking assessments, weekly reflective journals, and semi-structured interviews. Three customised agents: an AI Reading Assistant, an AI Speaking Partner, and an AI Speaking Examiner; were integrated into participants' learning routines. The findings revealed that GenAI contributed to improvements in pronunciation, grammatical accuracy, speaking fluency, and learners' confidence by offering immediate feedback and a low-anxiety learning environment. Participants also reported increased willingness to communicate and greater opportunities for repeated speaking practice in a supportive environment. In addition, GenAI-assisted interaction encouraged learners to engage in more self-directed speaking activities and facilitated greater awareness of their language problems during oral production. However, variations were observed across learners with different proficiency levels, particularly in the development of specific speaking micro-skills. Despite its pedagogical potential, GenAI also demonstrated limitations, including weak memory continuity, limited emotional authenticity, and difficulties handling culturally nuanced or complex communicative contexts. Overreliance on AI-generated responses during speaking practice may reduce opportunities for spontaneous language production and for independent use of communicative strategies. In addition, the accuracy and relevance of AI-generated feedback occasionally varied, leading to uncertainty among learners when evaluating and revising their spoken output. The results offer new insights into English-speaking instruction, emphasising the need for adaptive prompt engineering and teacher–AI collaboration. They also provide guidance on improving the use of GenAI in speaking instruction, such as designing scaffolded tasks for diverse proficiency levels.

Keywords: *AI agents, human-machine collaboration, English speaking, generative artificial intelligence*

Introduction

In recent years, the use of AI-driven English learning software has become increasingly widespread. Through simulated voice conversations, pronunciation assessment, and diverse feedback mechanisms, these systems provide learners with an efficient and personalised English-speaking learning experience, thereby enhancing their speaking skills. With the rapid evolution of artificial intelligence, Generative Artificial Intelligence (GenAI) further advances AI development and innovation in language education. Speaking is a key indicator of students' English proficiency.

However, traditional English-language instruction is hampered by a range of factors, including insufficient input, limited interaction, delayed feedback, and anxiety. GenAI technology offers new possibilities for addressing these challenges due to its localised adaptation, free and user-friendly nature, and efficient technical support, thereby becoming a vital tool in language learning. Although existing research has, to some extent, revealed the great potential of generative artificial intelligence in language learning, empirical research on GenAI-assisted instruction remains insufficient. Therefore, this study focuses on the domain of English speaking and conducts an empirical study with university English learners to deeply explore the effectiveness of generative AI in empowering English-speaking learning, thereby providing quantitative evidence for its application in this domain. The core research questions are: (1) How can GenAI-assisted personalised English-speaking learning be achieved? (2) What's the effectiveness of GenAI-assisted English-speaking learning on students' speaking proficiency?

Literature review

To provide a theoretical foundation for this empirical study, this literature review first discusses the key factors essential for developing English speaking ability and subsequently explores how Generative Artificial Intelligence (GenAI) is being leveraged to support personalised spoken language learning.

Key factors in English speaking

As a relatively complex cognitive process, speaking output in a foreign language is influenced by a combination of factors, including comprehensible input, interactive negotiation, linguistic feedback, and the affective filtering effect. Firstly, according to the input hypothesis in second language acquisition, a substantial amount of comprehensible input is a prerequisite for improving speaking ability. Therefore, students must continuously expand their vocabulary and sentence structures, refine their pronunciation, develop linguistic thinking, and cultivate a feel for the language by listening to and reading a wealth of high-quality linguistic materials that are slightly beyond their current proficiency level (Krashen, 1985). Secondly, the interaction hypothesis in second language acquisition posits that language learning is facilitated by the negotiation of interpersonal meaning and two-way interaction (Zhang et al., 2024). Constructivist learning theory also emphasises that students must actively construct their knowledge systems through collaborative interaction with others. Based on this, foreign language teachers need to strive to create authentic speaking interaction environments for students, guiding them in identifying speaking difficulties and receiving targeted language feedback during communication and meaning negotiation.

Furthermore, according to the affective filter hypothesis of second language acquisition, students' emotional factors, such as self-confidence, anxiety, and motivation, act as a filter for language input. Positive emotions facilitate the entry of language input into students' language acquisition mechanisms, while negative emotions hinder it (Krashen and Terrell, 1983; Zhang, 2024). Consequently, teachers should strive to create a relaxed and pleasant classroom atmosphere, design speaking activities that are appropriately challenging and engaging, reduce students' anxiety about speaking, and enhance their motivation and self-confidence.

GenAI empowers personalised English-speaking learning

In recent years, generative artificial intelligence technology has developed rapidly, demonstrating significant advantages in supporting students' spoken language practice. By functioning as both a 'language partner' and a 'feedback provider', it creates a relaxed, immersive language-learning environment, stimulating students' interest and initiative in spoken practice and thereby enabling truly open, intelligent spoken language learning. Scholars have conducted extensive research on the application of GenAI technology in foreign-language speaking practice. GenAI's powerful text comprehension and generation capabilities can comprehensively support language teaching, for example, by assisting teachers with lesson preparation, creating cross-cultural classrooms, and meeting the personalised learning needs of students at different proficiency levels. GenAI can provide a personalised, immersive learning experience for every language learner. Compared to traditional search engines, GenAI retrieval not only saves time but also provides highly relevant feedback and tailored learning recommendations, significantly promoting the development of thinking and language skills, and ensuring that learning content is highly aligned with the learner's characteristics. The chatbot Call Annie, developed using the GPT architecture, not only features an authentic English accent but also comes with a vivid animated avatar; it can also simulate speaking exams, provide speaking assessments, and offer practice recommendations (Wan and Moorhouse, 2025). These GenAI-driven automations of teaching and learning have reduced the burden on learners while enhancing learning efficiency (Jiao, 2023).

Furthermore, Dai et al. (2023) proposed a GenAI-based intelligent inquiry-learning model grounded in cognitive needs and question openness, aiming to enable progressive question-and-answer dialogues and iterative content generation through active questioning. English learners can pose a series of questions or requests in English, progressively seeking answers and resources from GenAI. The customised learning content and timely feedback provided by GenAI transform the traditional single-teacher instruction model into a multi-agent human-machine teaching model, thereby enabling intelligent production and delivery of specialised English knowledge (Wang et al., 2024). GenAI is capable of meeting learners' needs, objectives and learning habits. Thereby providing more avenues for personalised spoken English learning. This not only breaks the constraints of personalised learning being confined to the classroom, but also offers English learners richer channels for spoken input and platforms for output (Zou et al., 2025). However, empirical research on the use of GenAI to enhance spoken English learning remains in its early stages, and further empirical exploration is required regarding its specific potential, usage ways and attitudes towards its application in personalised spoken English learning.

Materials and Methods

This study collects data via a questionnaire and uses statistical analysis and semi-structured interviews to explore and interpret how Chinese university English learners achieve GenAI-enabled personalised speaking learning.

Research sample

This study took 48 non-English major undergraduates as participants in the survey. In the qualitative phase, 7 learners were voluntarily selected from among these 48 students to participate in semi-structured interviews.

Research tools

Doubao agent

The AI tool used in this study is Doubao, which offers multiple functions, including a chatbot, writing assistance, and English-learning support, and is available across various platforms. Based on the Doubao large language model, the teacher customised three agents to align with the teaching objectives and content: the “AI Reading Assistant”, the “AI Speaking Partner”, and the “AI Speaking Examiner”. These agents are trained on large amounts of professional data and possess extensive knowledge systems to serve various purposes and scenarios in English-speaking teaching.

Speaking test

The speaking test included a pre-test and a post-test. It assessed and diagnosed university students' speaking English proficiency based on the China Standards of English Language Ability. The test tasks included reading aloud, story retelling, picture description, picture-based commentary, speech delivery, and topic presentation.

Questionnaire

This study was designed to assess the “Questionnaire on the Effectiveness of Human-Machine Collaborative English Speaking Learning”. The first section of the questionnaire was designed to collect participants' background information. The second section incorporated elements of personalised learning. These included the types of commonly used GenAI tools, the frequency and duration of weekly use for foreign language learning, and the specific behavioural characteristics of students when using GenAI tools in their studies. After the questionnaire content was reviewed and evaluated by three university English teachers with extensive teaching experience, reliability and validity tests were conducted. The Cronbach's α coefficient of the questionnaire was 0.808, and the KMO value was 0.763, indicating high internal consistency and a reasonable factor structure.

Reflective journals and interviews

The learners wrote the reflective journals. After completing one week of practice, learners wrote a reflective journal entry that included (but was not limited to) a review of the week's speaking practice, an evaluation of the outcomes, and lessons learned. The interview was designed to investigate learners' feelings about using the Doubao agent for speaking practice. The interview consisted of six questions divided into four dimensions: perceived effectiveness, usability, satisfaction, and suggestions.

Research procedure

Phase 1: Preliminary preparation

This phase consisted of the following four steps. First, the researcher explained how to use the Doubao agents. Second, the researcher provided operational demonstrations.

Third, the learners tried out the Doubao agent and raised problems they encountered during use. Fourth, the researcher answered questions and clarified doubts.

Phase 2: A 30-Day speaking interactive practice using the Doubao agent

Before the practice began, the researcher administered a pre-test of the participants' speaking proficiency. Subsequently, participants were required to complete daily speaking-interaction tasks with the Doubao agents in their spare time and to send screenshots of the interaction interface to the researcher. The topics for the speaking interaction tasks were generated in three ways: designated by the teacher, jointly determined by the teacher and students, or chosen by the students themselves. The topics covered are friendship, views on money, views on romantic relationships, hobbies, and other issues closely related to the learners' daily lives. During the experiment, participants could contact the researcher at any time to analyse and resolve any confusion or problems encountered in their daily practice.

Phase 3: Post-test and interview

After all interactive practice tasks were completed, the participants took a post-test of their speaking proficiency and were interviewed by the researcher. After the experiment, the Questionnaire was distributed to investigate students' attitudes towards the human-machine collaborative teaching approach, especially its effects on students' speaking anxiety, self-confidence, interest, willingness to communicate, as well as the overall effectiveness of the teaching framework and their willingness to use GenAI for speaking learning in the future. After the questionnaire survey, in-depth interviews were conducted with 7 selected students to gain a detailed understanding of their personal experiences with the teaching framework and to verify the reliability of the experimental and questionnaire data from their perspective.

Framework for GenAI-Empowered English-Speaking Learning

Based on theories of second language acquisition and constructivist learning theory and drawing on AI-assisted English-speaking teaching methods proposed by relevant scholars, this study developed a GenAI-assisted English-speaking teaching framework. This framework set up multiple intelligent agents with diverse functions.

Technical preparation

Considering factors such as interaction design, response speed, user reviews, and applicability to English-speaking teaching, Doubao, which is used in China and has powerful chat interaction capabilities, was selected to support English-speaking learning. Based on the teaching objectives and content, three intelligent agents "AI Reading Assistant", "AI Speaking Partner", and "AI Speaking Examiner" were customised. These agents are trained on large amounts of professional data and possess extensive knowledge systems to serve various purposes and scenarios in English-speaking teaching. The role of the "AI Reading Assistant" agent is defined as a senior English-language expert who is a native English speaker. It possesses highly accurate speech recognition capabilities and a professional knowledge system of the English language. This enables it to comprehensively evaluate students' speaking in terms of pronunciation accuracy and fluency. Specifically, it can provide objective and fair

scores. It can also analyse students' errors in pronunciation, vocabulary, grammar, and content. Furthermore, it offers relevant improvement suggestions and provides correct pronunciation demonstrations. The role of the "AI Speaking Partner" agent focuses on guiding students in in-depth conversations based on textbook content, daily life topics, or simulated scenarios, patiently listening and encouraging students to speak, providing timely guidance without dominating the conversation, and correcting errors when appropriate, thereby helping improve students' fluency and accuracy in speaking English.

The "AI Speaking Examiner" agent can randomly generate speaking test tasks at an appropriate difficulty level, such as passage reading, individual presentations, and human-machine dialogues, and score students on pronunciation, expression, vocabulary, grammar, and content. Students can directly add these agents and set the language to English. In addition, with the assistance of the Doubao agent, teachers can guide students to master effective human-machine negotiation prompts in English, such as "Please correct my pronunciation" and "Can you read the vocabulary for me?", to ensure that different agents accurately understand and execute specific, actionable tasks.

Out-of-Class GenAI-assisted speaking practice

Students use GenAI outside of class for at least 20 minutes per day on speaking activities, such as answering questions, engaging in free dialogue, and negotiating with agents to retrieve information. These activities are designed to enhance their oral fluency, strategic competence, and real-time communication skills. First, using the AI Reading Assistant Agent, they read or repeat after textbook passages, vocabulary, class notes, exercises, and other classic English passages, and promptly correct errors using intelligent feedback and evaluation, especially for pronunciation tips. Second, using the AI Speaking Partner Agent, they engage in English conversations, including answering textbook-related speaking topics and conducting personalised human-machine dialogues based on their own English speaking proficiency and learning interests, for example, expressing opinions and engaging in in-depth discussions on interesting daily life topics, conducting simulated conversations in various scenarios such as business, travel, and socialising. Third, students use the AI Speaking Examiner Agent for speaking assessment activities, such as reading passages, individual presentations, and human-machine dialogues, at least once a week.

In-class GenAI-Assisted speaking practice

In the classroom, students use Doubao to engage in deep learning and speaking practice around the knowledge points of each unit. During the lead-in stage, the teacher organises students to use Doubao agents for brainstorming or small-group discussion speaking activities to activate their relevant background knowledge and stimulate learning interest. Teachers invite individual students to give live speaking presentations to check their out-of-class speaking learning outcomes. During the text-learning stage, the teacher organises students to use Doubao to read the text aloud, answer comprehension questions, or retell the content to digest the language and key points of the text. During the text summarising stage, the teacher organises students to engage in cooperative, project-based learning, such as role-plays, presentations, or debates, helping them creatively apply the knowledge learned and check their out-of-class speaking practice. In addition, the teacher provides targeted and personalised evaluation

and guidance on students' speaking performance, especially pronunciation accuracy, the appropriateness and complexity of vocabulary and grammar, and the completeness, creativity, and depth of content.

Results and Discussion

The following section presents the quantitative and qualitative findings of the study, detailing the effects of GenAI-assisted interactive practice on participants' speaking proficiency, their attitudes toward the approach, and their specific usage patterns.

Analysis of the effectiveness of GenAI-assisted speaking practice

Analysis of overall speaking proficiency

Based on participants' performance on the test tasks, a detailed assessment of their speaking proficiency was conducted. Comparing the pre-test and post-test diagnostic test results revealed that, except for Participant 6, whose speaking grade remained unchanged, all other participants achieved a higher post-test speaking grade. This finding indicates that after engaging in interactive speaking practice with the AI agent, most participants' overall speaking proficiency improved to some extent (*Figure 1*).

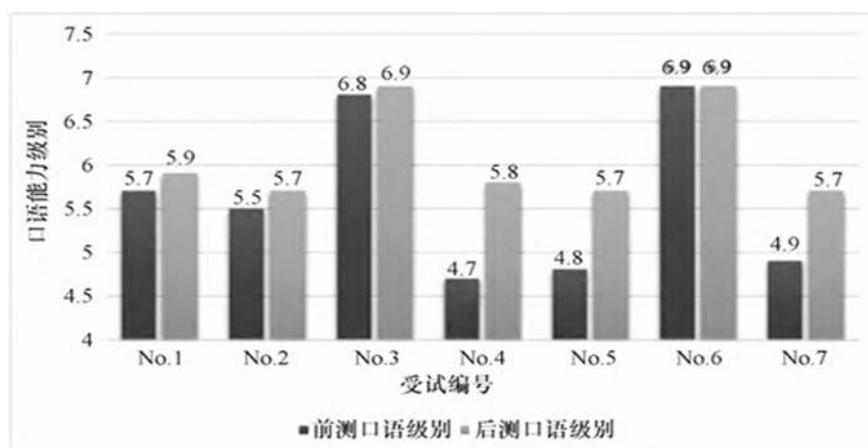


Figure 1. The effect of generative AI oral interactive practice on overall speaking proficiency.

Further analysis of the test results revealed differences in the extent of improvement among participants. Specifically, in the pre-test, three participants (Numbers 4, 5, and 7) were rated at the beginner level, two participants (Numbers 1 and 2) at the intermediate level, and another two participants (Numbers 3 and 6) at the advanced level. In the post-test, all beginner-level participants advanced successfully. Those at the Intermediate level all showed improvement. Among the advanced-level participants in the pre-test, one showed stable performance, with no change, while the other demonstrated slight progress (*Table 1*).

Table 1. Overall speaking proficiency performance of learners at different levels.

Proficiency Level	No.	Pre-test Score	Post-test Score
Beginner	4	4.7	5.8
Beginner	5	4.8	5.7

Beginner	7	4.9	5.7
Intermediate	1	5.7	5.9
Intermediate	2	5.5	5.7
Advanced	3	6.8	6.9
Advanced	6	6.9	6.9

Analysis of speaking micro-skills

Pronunciation

Pronunciation refers to the accuracy of the pronunciation of words and each syllable within a word. Among the seven participants, only one (Number 2) maintained the same pronunciation score without any improvement, while the pronunciation scores of the remaining participants all increased. Further observation of the pronunciation scores revealed that beginner learners (Numbers 4, 5, and 7) showed the most pronounced improvement, increasing by an average of 6 points. They were followed by intermediate learners (Numbers 1 and 2), who increased by an average of 2 points. The post-test pronunciation scores of advanced learners (Numbers 3 and 6) increased by an average of 1.5 points compared to their pre-test scores. This result indicates that learners at different proficiency levels exhibit different rates and patterns of pronunciation skill improvement, and that generative AI speaking interactive practice has the most significant effect on intermediate learners' pronunciation (*Figure 2*).

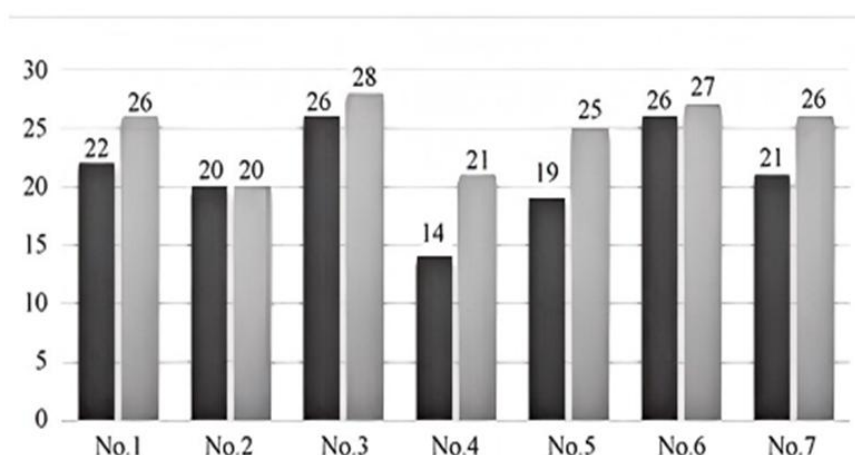


Figure 2. Effect of generative AI speaking interactive practice on pronunciation.

Language fluency

The evaluation criteria for language fluency involve the smoothness of expression, including the regulation of speech rate and the appropriate timing of pauses. Every participant made progress in language fluency, demonstrating clear improvement in both speech rate and the appropriate use of pauses. However, compared with the dimension of pronunciation, learners' progress in language fluency was relatively slower. Analysis and comparison of the pre-test and post-test fluency scores revealed that beginner learners (Numbers 4, 5, and 7) showed the most significant progress, with an improvement of 3.2 points. They were followed by intermediate learners (Numbers 1 and 2), with a 3-point improvement. Advanced learners (Numbers 3 and 6) showed the least progress, with an improvement of only 1 point. In the fluency dimension, the trend of improvement among high-, medium-, and low-proficiency learners was generally

consistent with that observed in the content dimension; that is, beginner learners showed the greatest improvement, followed by intermediate learners. In contrast, advanced learners showed relatively small improvement (*Figure 3*).

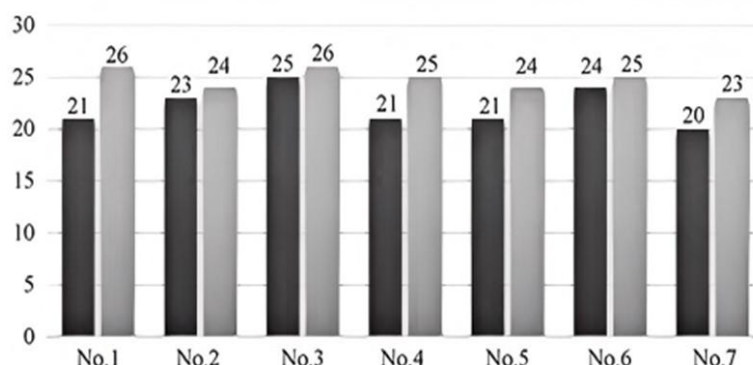


Figure 3. Effect of generative AI speaking interactive practice on speaking fluency.

Content

The scoring criteria for the content dimension cover the completeness of the reading task, such as whether the text was read in full, and whether there were any omissions or skipping, as well as performance on open-ended questions, including the clarity of key points, the completeness of details, and the adequacy of reasoning. A comparison of pre-and post-test data reveals that participants' performance in the content dimension improved across the board, as shown in *Figure 4*. Specifically, participants demonstrated greater accuracy and fluency in their reading of the material, whilst the logical coherence of their responses to open-ended questions also showed a marked improvement. This indicates that after practice, participants made substantial progress in content expression and organisational ability. It is worth noting that the rate and pattern of improvement demonstrated by learners at different proficiency levels in the content dimension align with those in language fluency, but differ from those in the phonetic dimension. Specifically, beginner learners (Numbers 4, 5, and 7) showed the most significant improvement in scores on the content dimension (5.3 points), followed by intermediate learners (Numbers 1 and 2), 2.5 points, whilst advanced learners (Numbers 3 and 6)' performance on this dimension remained relatively stable, with no marked improvement. This comparison reveals the complex relationship between proficiency levels and spoken micro-skills, highlighting the need to develop personalised teaching strategies.

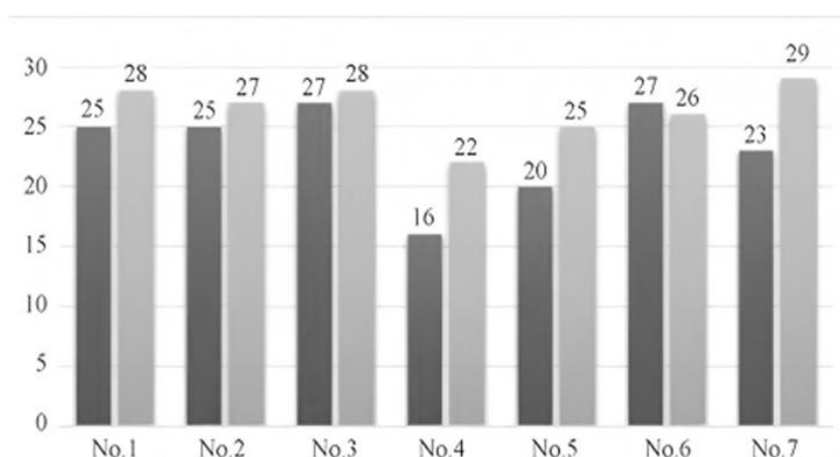


Figure 4. Effect of generative AI speaking interactive practice on speaking content.

Students' recognition of GenAI-assisted speaking practice

The statistical results of the 48 questionnaires collected from the experimental class are shown in Table 2. Overall, the average scores across all questionnaire dimensions were above 3.7, indicating that students held a relatively positive view of GenAI-assisted speaking practice. In the speaking-learning interest dimension, students believed that interacting with GenAI was close to real human communication, which greatly enhanced their interest in speaking English. In the dimension of speaking confidence, through diverse in-class and out-of-class speaking practice, students felt more confident when making individual presentations in class. In the dimension of speaking anxiety, GenAI created a safer environment for expression, and students reported feeling less nervous, even when making mistakes, during conversations with GenAI. In the dimension of willingness to communicate, the positive learning experience brought by GenAI encouraged students to engage more actively in communication. In the dimension of human-machine collaborative teaching effectiveness, personalised and precise evaluative feedback from GenAI improved the quality of students' spoken expression. In the dimension of intention to use GenAI, the practicality and effectiveness of GenAI motivated students to continue using it, and most students expressed willingness to continue using it for English-speaking practice in the future. In addition, students also highly recognised the GenAI-assisted speaking practice in the interviews. They indicated that GenAI was easy to use and served as an intelligent tool for speaking practice, effectively improving the efficiency of classroom speaking learning. Through a series of in-class and out-of-class speaking activities, they no longer felt as anxious about speaking English and became more confident, and ultimately improved their English-speaking ability.

Table 2. Survey results of EFL learners' attitudes towards GenAI-assisted English-speaking practice.

Dimension	Interest in speaking English Learning	Confidence in speaking and expression	Anxiety in speaking Communication	Willingness to Communicate	Human-AI Collaborative Teaching Effectiveness	Willingness to Use GenAI
Item No.	1, 2, 3	4, 5, 6	7, 8	9, 10, 11, 12	13, 14, 15, 16	17, 18, 19, 20
Mean	3.94	3.75	3.96	3.85	3.88	4.19

Students' usage patterns and behaviour characteristics

Tool selection

Regarding tool selection, students showed a clear preference for diversity. The teacher uniformly recommended the “Doubao” large language model and its three customised agents (AI Reading Assistant, AI Speaking Partner, and AI Speaking Examiner). The data showed that 91.7% of students used “Doubao” as their primary tool, while 8.3% frequently used other GenAI tools interchangeably, such as Kimi, WenxinYiyan, DeepSeek, etc. In addition, about 15% of students mentioned in interviews that they had occasionally tried ChatGPT or Qianwen.

Frequency and duration of use

In terms of usage frequency, 72.9% of students used GenAI tools for speaking practice 5 days or more per week, 18.8% used them 3-4 days per week, and only 8.3% used them 1-2 days per week. Regarding usage duration, 81.3% of students spent 20-30 minutes per session, 12.5% spent more than 30 minutes, and 6.2% spent less than 20 minutes. This distribution is largely consistent with the teacher's recommendation of “at least 20 minutes per day”, indicating that most students were able to carry out extracurricular autonomous AI speaking practice tasks conscientiously.

Usage activities

The questionnaire data categorised students' extracurricular AI usage behaviours into four types, with the participation proportions for each type as follows (*Table 3*).

Table 3. Students' extracurricular usage behaviors.

Activity Type	AI Agent Used	Specific Activities	Participation Rate
Reading Aloud & Shadowing	AI Reading Assistant	Repeat after textbook passages, vocabulary, and classic short essays; practice pronunciation repeatedly until scoring above 8 points	93.8%
Free Dialogue	AI Speaking Partner	Engage in free dialogue on daily life topics (e.g., campus life, travel, food, movies/TV); simulate business, travel, social and other scenarios	87.5%
Simulated Self-test	AI Speaking Examiner	Complete 1–2 full speaking simulation tests per week; review scores and feedback	81.3%
Information Inquiry	Doubao General Agent	Use spoken English to ask questions, find learning materials, translate sentences, and explain difficult grammar points	68.8%

Students demonstrated a high level of engagement and diverse speaking characteristics when using GenAI tools for extracurricular English-speaking practice. Regarding tool selection, although the teacher-recommended “Doubao” and its customised agents dominated usage, most students also actively tried other tools, such as Kimi and Wenxin Yiyan, indicating a degree of autonomous exploration. In terms of usage frequency and duration, most students met the daily practice requirements, and usage frequency was significantly and positively correlated with post-test scores, indicating that consistent practice is a key factor in improving speaking proficiency. With respect to specific speaking activities, reading aloud and shadowing, as well as free dialogue, were the most frequently adopted methods by students, followed by simulated self-tests and information inquiry. This reflects that students value both basic pronunciation training and communicative application in authentic contexts as well as self-assessment. Overall, students' GenAI usage behaviours exhibit a progressive

pattern, moving from structured practice to semi-open dialogue and then to autonomous inquiry, providing solid speaking data to support subsequent analysis of learning outcomes.

Results and Discussion

Based on theories of second language acquisition and constructivist learning theory, and drawing on AI-assisted English-speaking teaching methods proposed by relevant scholars, this study developed a GenAI-assisted English-speaking teaching framework. This framework set up multiple intelligent agents with diverse functions.

Positive effects of GenAI on speaking proficiency improvement

Research indicates that after engaging in interactive speaking practice with generative AI, most participants improved their speaking skills. This finding strongly supports the hypothesis of previous research that, with technological advancement, GenAI equipped with speech-interaction capabilities can accurately recognise users' voice input and rapidly generate fluent natural-language responses, thereby making it feasible to use GenAI to assist in speaking learning. Analysis of learners' interview data and reflective journals reveals that the positive effects of generative AI on English speaking are primarily evident in two areas.

High-quality Input and Instant Feedback Comprehensively Enhance Students' Speaking Proficiency

High-quality input

GenAI provides students with a large amount of high-quality, comprehensible, multimodal linguistic input, including speech, text, and knowledge, laying a solid foundation for improving English pronunciation, enhancing the accuracy and complexity of vocabulary and grammar in spoken expression, and deepening the content of expression. First, GenAI can produce high-quality reading demonstrations with precise pronunciation, natural and flexible intonation, and well-controlled rhythm for different contexts, providing models for students to imitate, thereby helping improve the accuracy of their English pronunciation. Second, GenAI can quickly convert human-computer interaction speech into text. By listening to and reading the dialogue content generated by GenAI, students are exposed to a large number of English expressions, with accurate vocabulary and grammar that conform to native-speaker habits. Over time, they can produce more accurate and complex vocabulary and sentence structures. Furthermore, GenAI possess excellent text comprehension and generation capabilities. Students can not only use them to expand their language repertoire but also increase their input of knowledge through continuous conversational negotiation, thereby making their spoken expressions more accurate and profound.

Instant feedback

By creating an all-time, multi-scenario interactive environment, GenAI provides personalised dialogue guidance and precise pronunciation feedback, helping students adjust and optimise their language expression during human-computer interaction, achieve deep intellectual engagement, and comprehensively enhance their ability to

express themselves in speech. GenAI creates a relaxed and pleasant, immersive, and interactive environment for students. They can engage in various speaking dialogue practices anytime, anywhere, such as daily conversations, scenario-based simulations, and discussions of textbook topics. Extensive human-computer interactive speaking negotiation helps students actively construct and improve their speaking expression, making their speech more fluent and smoother. Relying on a vast knowledge base, GenAI can deeply analyse the semantic meaning, grammatical structure, and contextual logical relationships in students' spoken dialogues, quickly generate natural, smooth real-time responses, and provide encouraging feedback when students experience dialogue stagnation, awkward silences, or topic deviation. They generate prompts or new open-ended questions to inspire and guide students to expand their arguments, sustain the dialogue, and thereby enhance their speaking fluency. In addition, using advanced speech recognition and assessment technologies, standard English pronunciation rules, and large amounts of linguistic data, GenAI can help students identify speech errors and automatically correct them. That is, they can objectively and accurately locate and analyse students' pronunciation, intonation, speech rate, and other phonetic features, provide relatively scientific, comprehensive scores, and simultaneously offer corrective suggestions and standard model pronunciations for students' pronunciation errors, enabling students to revise and perfect their pronunciation and construct a correct phonological system.

In this study, the Doubao agent not only focused on learners' speaking output but also provided them with feedback on content and structure, as well as fine-grained skill feedback at the linguistic level. Behaviourist psychology holds that feedback serves as a bridge between input and output. Effective feedback can promptly adjust learners' learning behaviours and facilitate the achievement of learning goals. The feedback provided by generative AI is both efficient and interactive. On the one hand, it can quickly identify and correct students' errors and provide correct expressions. On the other hand, the diverse vocabulary and sentence patterns used by generative AI serve as models for learners to study and imitate, enriching their language expression and thereby effectively improving speaking proficiency. This was also verified in interviews and reflective journals. For example, Participant 5 noted in a reflective journal: "Doubao corrects some of my pronunciations, such as when I forget to add 's' to a word for plurals or third-person singular verbs, and some words need to be in the past tense. During the conversation, if there are grammatical errors, it corrects me promptly." Participant 2 pointed out in an interview: "I think one aspect that helps me a lot is that it first summarises my answers in more refined and advanced language." During this summarisation process, I have the opportunity to learn its way of expressing things and then apply it in subsequent conversations."

Enhancing students' interest and confidence in speaking

GenAI helps enhance students' interest in learning, build confidence in expressing themselves, alleviate communication anxiety, and ultimately strengthen their willingness to communicate. First, GenAI's speech feedback is similar to that of real humans, enabling students to experience a sense of achievement and fun during interactions, like real communication. For example, the participating students in this study reported that they could freely begin their speaking practice based on their own English proficiency and learning interests, gaining greater autonomy and a sense of control over setting learning goals and managing the learning process, while gradually

building interest and confidence in speaking. GenAI breaks the boundaries of time and space, interacting with learners naturally and positively. Whenever and wherever they are, learners can engage in seamless communication with GenAI. During the interaction, the GenAI's pronunciation is lifelike, almost indistinguishable from that of real people. Learners can also choose voices of different genders or styles according to their preferences, making the communication experience closer to real interpersonal interaction and thereby giving students a sense of achievement and enjoyment.

Secondly, compared with face-to-face communication with real people, students experience less pressure regarding their social image when interacting with GenAI. Moreover, GenAI is extremely patient during conversations with users, answering every question and providing positive feedback and encouragement, effectively bridging the psychological distance with students. Therefore, students reported feeling very relaxed when chatting in English with GenAI and did not feel shy or anxious even when making mistakes in expression. The non-human nature of GenAI effectively alleviates learners' tension, encourages them to speak and express themselves, and thus promotes improvements in learners' speaking proficiency. When asked about their experience using the Doubao agent, learners frequently mentioned keywords such as "positive", "companionship", "emotion" and "help". For example, Participant 3 shared: "The Doubao agent is an excellent listener. No matter what topic I discuss with it, it always gives me positive feedback. It accompanies me like a friend, which not only makes me more courageous in expressing myself in English and boosts my confidence, but also encourages me to express my emotions and ideas bravely."

Furthermore, the participating students in this study reported that by persistently chatting in English with GenAI, they became more confident in their English-speaking expression. Whether called upon to answer questions or asked to perform role-plays or speeches on stage, they no longer felt as nervous or anxious. These positive emotional factors prompted them to be more willing to participate actively in classroom speaking activities, to look forward to more opportunities for English practice outside class, and even to proactively expand their communication channels, such as joining English clubs or engaging in English conversations with international students on campus.

Different learner levels exhibit different performances in speaking micro-skills

The study found that after GenAI -assisted speaking interaction, most participants significantly enhanced their speaking abilities. However, learners at different proficiency levels showed marked differences across various dimensions of speaking performance. Specifically, in the pronunciation dimension, intermediate learners made the most prominent progress; in the content and fluency dimensions, beginner learners showed the most significant improvement. In the pronunciation dimension, intermediate learners made notable progress. This phenomenon may be closely related to their being at a critical transitional stage in the development of speaking skills. At this stage, learners' phonological systems are highly open and plastic, enabling them to more easily absorb and imitate new pronunciation patterns. During interviews, when asked, "What gains have you made after one month of speaking practice?" Intermediate learners generally reported that their pronunciation had significantly improved through interaction with the Doubao agent. For example, Participant 1 shared: "I consciously imitated Doubao's pronunciation, and as a result, I found that my pronunciation not only became more accurate but also sounded more natural and fluent. Doubao's

feedback was very timely; whenever I made a pronunciation error, it pointed it out immediately, allowing me to correct the mistake on the spot.”

In the content and fluency dimensions, beginner learners tended to show more substantial progress than intermediate and advanced learners. This difference relates to how learners allocate attentional resources. During language processing, there is a competitive relationship between content and form (Skehan, 1998). Typically, language learners prioritise the content of language, and this phenomenon is particularly pronounced among beginner learners. Beginner learners tend to focus on the meaning conveyed by language rather than excessively worrying about linguistic accuracy, in stark contrast to intermediate and advanced learners. Intermediate and advanced learners need to attend to both linguistic form and meaning simultaneously, and this multitasking requires sacrificing quality in one to improve the other. This was fully evident in the dialogue transcript between Participant 3 and Doubao. During the conversation, Participant 3 frequently engaged in self-correction, which repeatedly interrupted his content expression. In a subsequent interview, he also mentioned that he could not help thinking about whether his word choices were accurate or sophisticated enough, and that he often forgot the next point he wanted to make and experienced speech breakdowns.

The differences in the rate and pattern of speaking-microskill improvement among learners at different levels provide important insights for formulating subsequent speaking-teaching strategies. First, when using GenAI for speaking instruction, teachers must fully consider learners' characteristics at different levels and design targeted speaking instruction interventions. For intermediate and advanced learners, instruction should focus on expanding the depth and breadth of their vocabulary knowledge and enhancing the diversity of their language expression. For example, when creating an agent, teachers should configure it as an advanced speaking partner, requiring it to provide precise vocabulary explanations and usage examples during speaking interaction, and introduce more complex, contextually appropriate vocabulary and expressions. For beginner learners, the instructional focus should be on increasing the frequency and automaticity of speaking output, as well as enhancing the precision of pronunciation training. When configuring the agent's role, the emphasis should be on fluent content expression. By providing a “friendly” speaking partner, beginner learners can practice repeatedly and gradually establish an automatic mechanism for language output, making their expression increasingly fluent and natural.

Certain limitations of generative AI in promoting speaking and learning

Although GenAI significantly enhances learners' speaking learning outcomes, it still has certain limitations. Through interviews with learners and analysis of the interview results, we found that these limitations are primarily manifested in three areas: insufficient memory continuity, a lack of emotional communication, and inadequate handling of complex topics. Memory continuity refers to the system's ability to retain previous dialogue content and context, which is crucial for maintaining natural, smooth conversations. In real interpersonal communication, we rely on the other person's memory to maintain topic coherence, understand each other's intentions, and construct in-depth dialogue. However, GenAI has limited capabilities in this regard. It is generally unable to remember details of previous conversations, leading to a lack of coherence in dialogue and making it difficult for learners to experience an authentic communication environment. For example, one participant (No. 7) mentioned: “It lacks continuity. It's

not like a real person with memory who can engage in in-depth communication with you. I find this quite unsatisfactory.”

Emotional communication is at the core of interpersonal interaction, involving the ability to express and understand emotions. Generative AI falls short in simulating emotional communication. It cannot recognise and respond to emotional nuances and cannot provide empathetic feedback. This limits learners' ability to improve their emotional language expression and comprehension, making interactions seem rigid and lacking in resonance. When asked about the shortcomings of GenAI, Participant No. 6 mentioned: “Although it always stays with you, it is an AI without its own opinions. It always supports your views. Moreover, when I tell it that I'm feeling down, its comforting words follow the same pattern. It also cannot joke with you like a real person.” Furthermore, when discussing sensitive or complex topics, GenAI often performs unsatisfactorily, avoiding them and failing to provide appropriate responses. This approach not only limits learners' opportunities to explore and practice discussing such topics but also hinders their ability to improve language expression and topic comprehension in real-world contexts.

Conclusion

This study systematically validated the utility of Generative AI technology, implemented through a novel three-agent framework on the Doubao platform, for enhancing language learners' English-speaking proficiency. The findings offer empirical evidence and theoretical nuance regarding GenAI's differential impact on micro-skills across proficiency levels and are as follows: (1) Generative AI has a positive effect on driving the improvement of speaking proficiency, significantly enhancing learners' overall speaking ability, pronunciation accuracy, richness of language content, and fluency. This positive effect mainly stems from the immediate feedback and emotional support provided by GenAI. The former helps learners adjust their learning behaviours promptly, while the latter strengthens learners' motivation and self-confidence. (2) Learners at different proficiency levels exhibit differences in their speaking micro-skills. Specifically, intermediate learners showed the most pronounced progress in pronunciation, whereas beginner learners demonstrated greater improvement in content and fluency. This variation relates to how learners allocate attentional resources across different stages of learning. This finding has important implications for formulating subsequent speaking-teaching strategies: teachers should adopt differentiated instructional interventions for learners at different levels to maximise the potential of GenAI in speaking learning. (3) Generative AI still has certain limitations in speaking practice. Insufficient memory continuity, lack of emotional communication, and inadequate handling of complex topics are the main issues currently facing. These problems constrain learners' language experience and the improvement of their comprehension abilities in authentic communication contexts and affect the comprehensive application of GenAI in speaking learning. Therefore, when using GenAI for speaking instruction, teachers should leverage its strengths while addressing its potential limitations, thereby optimising the effectiveness of speaking instruction and promoting the quality and upgrading of English instruction.

GenAI is a highly effective, operational intelligent teaching tool that offers significant advantages in improving students' speaking ability, willingness to communicate, and classroom effectiveness. Therefore, EFL teachers should actively

integrate it into instruction to promote efficient and personalised learning through human-machine collaboration. However, its implementation requires attention to two critical aspects: teachers must guide students in critically evaluating GenAI-generated content to prevent overreliance, and they must address GenAI's limitations in non-linguistic modalities (such as facial expressions and gestures). Since spoken communication is multimodal, cultivating students' ability to coordinate linguistic and non-linguistic features is essential for effective and expressive communication. While this preliminary exploration successfully verified the positive role of GenAI through mixed methods, its limitations, i.e. notably the small sample size and the assessment tool's generalisability constraints, guide the direction of future scholarship. Future research must expand the participant pool and refine assessment methods to fully capture GenAI's instructional effectiveness. Ultimately, this work reaffirms that GenAI, through human-machine collaboration, is not merely a supplementary tool but a transformative pathway essential for the efficient, personalised, and upgraded future of English-speaking instruction.

Acknowledgement

This research is self-funded.

Conflict of interest

The authors confirm that there is no conflict of interest among the parties involved in this research study.

REFERENCES

- [1] Krashen, S.D. (1985): The input hypothesis: Issues and implications. – Longman, London 120p.
- [2] Krashen, S.D., Terrell, T.D. (1983): The natural approach: Language acquisition in the classroom. – Pergamon Press, Oxford 191p.
- [3] Dai, L., Zhao, X.W., Zhu, Z.T. (2023): Smart inquiry: A new mode of conversational learning based on ChatGPT. – Open Education Research 29(6): 42-51.
- [4] Jiao, J.L. (2023): ChatGPT boosting the digital transformation of school education: What to learn and how to teach in the age of artificial intelligence. – Chinese Journal of Distance Education 43(4): 16-23.
- [5] Wan, Y., Moorhouse, B.L. (2025): Using Call Annie as a generative artificial intelligence speaking partner for language learners. – RELC Journal 56(2): 489-498.
- [6] Skehan, P. (1998): A cognitive approach to language learning. – Oxford University Press, Oxford 324p.
- [7] Wang, C., Wang, Y., Zou, B. (2024): Revolutionising EFL pedagogy: Innovative strategies for integrating GAI (ChatGPT) into language teaching. – Journal of Language Teaching 4(1): 1-7.
- [8] Zhang, C., Meng, Y., Ma, X. (2024): Artificial intelligence in EFL speaking: Impact on enjoyment, anxiety, and willingness to communicate. – System 121: 35p.
- [9] 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.

- [10] Zou, B., Wang, C., Yan, Y., Du, X., Ji, Y. (2025): Exploring English as a foreign language learners' adoption and utilisation of ChatGPT for speaking practice through an extended technology acceptance model. – International Journal of Applied Linguistics 35(2): 689-704.