



## From silence to speech: Enhancing student oral communication skills through AI-Chatbots

Saman Ranabahu<sup>1</sup>

<sup>1</sup> ELICOS Teacher, Monash College, Australia

Ram Ashish Giri<sup>2</sup>

<sup>2</sup> ELICOS Teacher, Monash College, Australia

Alastair James<sup>3</sup>

<sup>3</sup> Programme Leader, English, Monash College, Australia

### Abstract

This article explores how an AI-powered chatbot (Doubao) could support the development of English oral communication skills among 14 postgraduate students enrolled in an English for Academic Purposes (EAP) course, in an English Language Intensive Courses for Overseas Students (ELICOS) context. The students initially engaged in structured interactions with Doubao to build familiarity and speaking confidence. These activities were then integrated into course assessment preparation, focusing on tutorial discussion and oral presentation skills. Data collection included results from the diagnostic and final speaking assessments, student audio reflections, focus group interviews, and teacher class notes. Findings revealed that the use of Doubao enhanced learners' fluency, confidence, and awareness of communication strategies, while providing a space for practice and feedback. The study highlights AI's potential as both a pedagogical aid and a collaborative language partner, demonstrating that chatbot-mediated practice can effectively complement classroom instruction and promote learner autonomy in EAP and ELICOS settings, while also acknowledging that a lack of human connection remains a barrier for some students.

### Correspondence

Saman Ranabahu  
[saman.ranabahu@monashcollege.edu.au](mailto:saman.ranabahu@monashcollege.edu.au)

### Publication

Received: 08 December 2025  
Revision: 10 April 2026  
Accepted: 17 April 2026  
Published: 30 June 2026

### Copyright

©2026 TESOL in Context



This work is licensed under a [Creative Commons Attribution – ShareAlike 4.0 International License](https://creativecommons.org/licenses/by-sa/4.0/).

**Keywords:** *AI-assisted learning; oral skills; learner engagement; tutorial discussion skills; oral reflection; presentation skills.*

## **Introduction**

Developing oral communication skills is a persistent challenge for learners in pathway English courses, particularly at B1–B2 (Intermediate – Upper Intermediate) levels, where students are expected to engage in academic discussions and presentations but often show reluctance to speak. In this context, EAP courses aim to prepare postgraduate students for the communicative demands of higher education, yet many learners struggle with confidence, fluency, and appropriate use of communication strategies. This study explores how an AI-powered chatbot, Doubao, can support these learners by providing structured opportunities for speaking practice. Fourteen postgraduate students enrolled in an EAP course based in Suzhou, China, and delivered by an Australian institute, initially engaged in guided interactions with Doubao to build familiarity and confidence. The tool was later integrated into preparation for tutorial discussions and oral presentations.

## **Literature review**

The existing body of research on the use of chatbots and AI-powered tools for developing speaking skills has been primarily situated within English as a Second Language (ESL), English as a Foreign Language (EFL) and ELICOS contexts. These environments cater to learners pursuing English with the aim of attaining high levels of proficiency, particularly for academic or professional purposes. As Nguyen (2025) observes, much of the existing research on AI-assisted speaking practice centres on university-level EFL learners, reflecting the global importance of English for communication, science, and media in an increasingly interconnected world.

A number of studies have explored the effects of AI tools across diverse learner populations within higher education and language programs. Al-husban (2025) and Alwin et al. (2024), for example, demonstrated that the English Language Speech Assistant (ELSA) speech analyser significantly enhanced university EFL students' speaking skills across multiple criteria, pronunciation, intonation, fluency, vocabulary, and grammar, compared to traditional teaching methods. In the context of adult education, Chen et al. (2022) examined the use of AI tools in a workplace based English Language program, which were designed to improve communicative ability through authentic speaking tasks that mirror real-life language use situations. Other studies have targeted vocational students, showing that tools such as Mission Fluent, an AI chatbot, were both user-friendly and effective in improving pronunciation (Hoang et al., 2023). Similarly, AI-based mobile applications have been shown to support Saudi Preparatory-Year students in developing EFL speaking proficiency (Makhlouf, 2021). A systematic review of 24 studies (2017–2023) by Du and Daniel (2024) found that research on chatbots for English-speaking practice is still developing, while current evidence shows that AI chatbots can speed up learning, reduce speaking anxiety, improve pronunciation, and enhance students' confidence, engagement, motivation, and overall speaking outcomes.

Within ELICOS and other intensive academic contexts, AI tools have been recognized for addressing the linguistic demands of higher education preparation. Many projects bring together applied linguistics and data science tailoring AI chatbots to help students in ELICOS to practice real-life communication using both text and voice (Lee et. al., 2025). They sought to offer practical, evidence-based methods for incorporating conversational AI into language learning programs on a larger scale. Such programs created language models (Singh & Namin, 2025) for students to use in their speaking practice. For more advanced learners, Zhu et. al. (2025) argue that AI-assisted mobile applications can facilitate mastery of complex academic speaking skills; critical thinking, argumentation, and precision, which are required for English for Academic Purposes (EAP). El Shazly (2021) further highlights positive student attitudes towards these tools, suggesting that they enhance motivation and self-directed learning. Mohamed (2024) stresses that the validity of AI-assisted assessment depends on whether it authentically reflects the communicative and academic demands of English-medium universities. Mohammadkarimi (2024) and Qiao and Zhao, (2023) also emphasize that such systems offer detailed feedback on pronunciation, fluency, grammar, and vocabulary, enabling students to identify weaknesses and make measurable progress. Nguyen (2024) reinforces this view, noting that AI-driven feedback effectively fills learning gaps left by traditional instruction, thereby promoting autonomous learning.

Pedagogically, integrating AI technologies aligns with broader institutional goals and teaching efficiencies. As Pokrivcakova (2019) observes, universities increasingly incorporate AI-assisted language learning to produce fluent, employable graduates. From a teaching perspective, Han (2020) suggests that students interacting with AI Chatbots improved in speaking ability, and their motivation, interest, and perceived anxiety changed positively compared to before using the chatbot. Additionally, Yusuf et. al. (2025) found that AI-based assessment platforms reduce the subjectivity and workload associated with speaking evaluations, helping teachers focus on innovative instructional design. Furthermore, in high-pressure intensive programs, AI chatbots can mitigate learners' anxiety by providing a stress-free environment for practice. Tai and Chen (2024), Jeon et. al (2023) and Zou et al. (2020) highlight those students who practice speaking with AI interlocutors often report reduced fear of making mistakes and increased confidence in spontaneous communication.

Regarding the AI-powered chatbots, such as Doubao, a growing body of research highlights the potential of them to address persistent challenges in developing English for Academic Purposes (EAP) speaking skills. Ma et al. (2024) found that consistent interaction with Doubao significantly reduced Foreign Language Anxiety among Chinese EFL undergraduates by providing a low-pressure, non-judgmental environment for practice. Building on this, Liu et al. (2025) demonstrated that an AI-blended approach incorporating Doubao enhanced learner self-efficacy, goal orientation, and academic performance in complex skills like listening and writing. Similarly, Shikun et al. (2024) confirmed that students using chatbots like Doubao improved pronunciation, intonation, and confidence, though they noted limitations such as imperfect speech recognition and restricted topic ranges. Collectively, these studies suggest that chatbots like Doubao can create supportive

environments that reduce anxiety, build confidence, and improve both foundational and academic speaking skills in EAP contexts.

Overall, the literature substantiates the value of chatbots and AI speech evaluation tools in structured ESL and EFL programs, particularly those comparable to ELICOS contexts. Research consistently demonstrates that these technologies not only enhance learners' fluency, accuracy, and pronunciation but also address academic and affective dimensions of language learning. As Nguyen (2025) concludes, such tools play a vital role in achieving the high-level speaking proficiency required for academic and professional success, underscoring their growing relevance in modern language education.

### ***Research gaps***

Despite emerging evidence that AI chatbots can enhance speaking skills by reducing anxiety and improving pronunciation and fluency (Ma et al., 2024; Shikun et al., 2024), significant research gaps remain regarding the efficacy and limitations of this technology. Firstly, it is unclear whether these tools can effectively develop higher-order EAP speaking competencies, such as argument construction, critical response, and academic register appropriateness, beyond basic oral fluency (Liu et al., 2025). Secondly, the challenges students face when using AI chatbots for speaking practice are underexplored. The available literature preliminarily indicates technical limitations, including imperfect speech recognition and restricted and repetitive interactions, which can lead to disengagement and frustration (Shikun et al., 2024). Furthermore, the quality and pedagogical appropriateness of corrective feedback provided by current AI tools remains questionable, as students report disappointment when errors go unrecognised or uncorrected (Shikun et al., 2024). Addressing these gaps is essential to determine whether AI chatbots can genuinely support the complex communicative demands of academic ELICOS contexts and to identify the pedagogical scaffolds needed to mitigate their current limitations. With the view to understanding the potential of AI chatbots, the present study attempts to address the following research questions.

### ***Research questions***

RQ1: To what extent can generative AI, Doubao, enhance ELICOS students' speaking skills?  
RQ2: What challenges do students face while using chatbots such as Doubao to enhance their oral skills?

## **The research design**

### ***Participant selection***

Eighteen participants in the first phase of the study were conveniently selected from one of the researchers' classes in an English Module 4 (pre-advanced, CEFR B2-C1) in a Melbourne ELICOS institute. The purpose of this phase is to see which chatbot students preferred to use for practicing speaking skills. Subsequently, participants for phases 2 and 3 were purposefully selected from two of the twenty English Degree Programme (EDP) classes, at similar language levels. This program was delivered in an institute based in Suzhou, China. In total, 14 students were chosen from the researchers' own EDP classes. The EDP is a customised EAP pathway supporting a joint Australia-China postgraduate program delivered collaboratively by two leading universities - one in Australia and one in China. The participants, aged between 22 and 26, were first preparing English competency in the EDP program in China, to pursue master's degrees later in disciplines such as engineering, health, translation and interpretation, and information technology.

Data for this project were generated from two groups: experimental and control groups. The experimental group consisted of 14 students from two classes (seven students from each class). These students were directly consulted by the researcher/ teachers, and this academic consultation last between 10 and 15 minutes each, and was conducted on a weekly basis. Prior to the study, all experimental group participants attended a short orientation session in which the purpose and procedures of the research were clearly explained. The control group included 14 students who were selected randomly from the rest of the two class lists (seven students from each class). These students were not directly consulted by the researchers/ teachers. Having an experimental group and a control group enabled the researchers to compare their speaking performance and determine whether the use of the Doubao app led to measurable improvements in speaking skills.

### ***Data collection***

Data was collected from 28 students (14 students from experimental groups and 14 from the control groups). The orientation was provided to the experimental group only as they needed to understand the process of trialling the Doubao app. The scores from the control group were collected from their assessment records (Tutorial Discussion and Spoken Reflection Presentation) with the permission from the programme leader.

Data were collected using multiple sources: student scores from the Week 1 diagnostic speaking task, the Week 6 tutorial discussion, and the Week 10 spoken reflection assessment, as well as weekly student reflection audio recordings. In addition, a focus group discussion was conducted in Week 10 at the conclusion of the data collection process. During the

tutorial discussion and spoken reflection assessment, teacher class notes were recorded by the researchers/ teachers.

### *Ethical considerations*

Prior to data collection, the researchers provided 28 participants with a thorough explanation of the study's procedures, expectations, and ethical standards. Participants were informed about their rights, including voluntary participation and confidentiality. After these explanations, they signed written consent forms confirming their willingness to take part in the project. Following discussions about potential tools, participants collectively agreed to use the Doubao app as the main platform for practising English speaking skills during the study period. The students were able to use the free version of the app without any issues for the purposes of the study. The control group students were not required to submit their consent as they did not participate in the trial.

Research participants and researcher's power relationship is a significant issue in qualitative classroom-based research, and our research is no exception to this. According to Kernielli-Miller and Heifer (2019), qualitative research aspires to redistribute power between researchers and participants, but in practice the researcher still holds considerable control over the process. This tension is especially sharp when participants are students, because they may feel obliged to cooperate when the researcher simultaneously occupies the role of their teacher. Even when the interactions treat them as "equals" during meetings and/or interviews, the researchers decide how to analyse, edit and represent what has been said, which can recreate hierarchy despite an egalitarian ethos.

This study, therefore, stresses the need for explicit clarification of roles, rights and limits at each stage of the research. The research team consisted of two teachers who conducted the trial, along with a third member, the programme lead, who oversaw and supported the project. The researchers provided their ethical commitment to no-harm and students' welfare over their own academic goals by paying careful attention to informed consent, giving opportunities to participants to question, challenge or even withdraw from the research. In addition, the researchers sought assistance from a local teacher in China to explain the project processes using the participants' first language highlighting their rights of voluntary participation and their right to withdraw at any stage of the project.

The power relationship between the participants and researchers became more balanced as leadership emerged from within the participating students. One participating student assumed the role of team leader for both groups and coordinated communication between the researchers and other participants. The team leaders also coordinated the Doubao trialling and the submission of reflection recordings. Furthermore, the team leaders served as the first point of contact for participants who encountered difficulties, such as understanding the task prompt or retrieving their recordings from their devices.

### ***Three main phases of study***

The study was completed in three iterative phases, and each informed and shaped the subsequent phases.

#### *Phase 1.*

The first phase of the study, which can be described as an informal research tools selection phase, was conducted in Melbourne to identify the most appropriate tool for use on the main study. This phase focused on tool selection and was informed by student feedback following the researcher's initial framing. This planning phase began with general inquiries into which apps students commonly used to develop and practice their speaking skills. This quickly evolved into an organic debate among students, who compared and discussed the merits of different speaking skills learning applications. As the process unfolded, students actively demonstrated the apps they used and explained their personal approaches to using them effectively. The researcher observed these demonstrations and took notes on the students' interactions, comments, and emerging preferences. Through this observation, it became evident that Doubao was the favoured application among participants and was therefore selected as the research tool. All participants were Chinese-background postgraduate students at their respective universities and had prior experience with the tool.

#### *Phase 2.*

After completing the first phase with the students in Melbourne, the researching teachers travelled to Suzhou, China to teach the EDP course. It is during the EDP course the second and third phases of the research were conducted. The goal of the second phase of the study was to familiarise both the students and the participating teachers with the use of the chatbot Doubao. This stage, which lasted four weeks, proved pivotal, as the questions, experiences and familiarities developed during this process shaped the subsequent phases. As the project progressed, students gradually gained confidence in interacting with the chatbot, which not only improved the quality of research outputs, but also heightened student engagement, particularly as the activity became tied to the course's assessment criteria in phase 3.

The researchers initiated this phase 2 by organising an orientation and a demonstration session. The participants from the two different classes were brought together in a meeting room where the researchers outlined the overall task, and explained the process and required activities. With support from a local colleague, whose presence helped demonstrate effective use of the app, provide the study with greater legitimacy in students' eyes, and foster a sense of belonging, the first chatbot prompt was introduced and demonstrated. Students were then instructed to engage with the chatbot five times a week and produce four audio reflections, one for each of the four weeks of the phase. The initial prompt was intentionally broad, asking students to discuss their familiarity with the tool, their experiences after using it, and their thoughts and feelings about their weekly learning interactions.

As the study progressed, the researchers observed that students from the two classes began operating as a single, cohesive team. A team leader emerged within the group, naturally taking on the role of liaison between the researchers and the participants. Though the students initially came from two distinct classes, their shared experience allowed the researchers to analyse their progress as one unified cohort. In reflection, the researchers noted that this spontaneous team dynamic became an important catalyst for their deeper engagement. The informal peer support system encouraged students to speak openly about their growing confidence, as well as the confusions, constraints, and limitations they encountered throughout their journey.

### *Phase 3.*

In the third and final phase of the project, shaped by lessons and experiences from the earlier stages, the focus shifted to the assessment phase of the EDP course, targeting core elements of two speaking assessments: Tutorial Discussion and Spoken Reflection. Over the last four weeks, students worked with two tailored prompts aimed at improving their discussion techniques and presentation skills. Their interactions with the chatbot, Doubao, ranged from asking it to clarify complex academic concepts such as “acculturation” and “deculturation”, to learning conversational strategies like interrupting politely and managing turn-taking. For the spoken reflection assessment, students used Doubao to rehearse potential audience questions and formulate effective responses, turning the tool into both a rehearsal partner and a source of strategic guidance.

At the conclusion of the study, the researchers conducted a semi-structured focus group interview involving both groups of 28 participants. The session lasted approximately 45 minutes, during which participants shared their experiences and reflections. The interview was audio recorded, and additional teacher notes were taken to capture non-verbal observations and contextual insights.

### ***Data triangulation***

The triangulation of the data/findings was established through several strategies. Credibility was ensured by employing multiple data sources like triangulation, conducting regular peer debriefing sessions among the three researchers before and throughout the study, and implementing successive trials across two distinct courses (ME4 in Melbourne and EDP in China) with different cohorts of students to achieve relatively prolonged engagement. Regarding transferability, participants were purposefully selected to generate rich, relevant data closely aligned with the research question. Dependability was maintained by keeping comprehensive records of the research process to facilitate potential replication and by engaging in regular peer discussions to promote methodological consistency and rigor. The positionality of the researchers is also acknowledged, as they share the same institutional setting and operate within the same social and cultural context as the study. This shared

background may have contributed to a convergence of perspectives and professional commitments.

### ***Data collection and analysis***

Qualitative data were sourced from the weekly reflection audios, a focus group interview, and teacher class notes. Meanwhile, quantitative data were drawn from participants' diagnostic speaking assessments and final speaking assessments, which included a tutorial discussion in week 6 of the program and the spoken reflection presentation in week 9. These assessments provided measurable indicators of progress in speaking performance. In addition to recorded scores, researchers kept detailed class observation notes aligned with the assessment rubrics to support a comprehensive analysis of both qualitative (from weekly reflection audio files, the focus group interview, and teacher notes), and quantitative findings.

The analysis of the data involved both qualitative and quantitative approaches. For the qualitative component, weekly reflection audio (WR audio) files from the experimental groups were transcribed, along with transcriptions from a semi-structured focus group interview conducted at the conclusion of the study. All transcripts were uploaded to the data analysis platform Taguette, where the information underwent systematic coding. Once coded, key themes were extracted, organised into tables, and interpreted in alignment with the study's guiding research questions. Quantitative data sets were tabulated and carefully analysed to reinforce and complement the findings emerging from the qualitative analysis.

## **Findings**

The quantitative and qualitative data both reveal that Generative AI Doubao enhanced EFL / ESL students' speaking skills in more than one way. Its interactive nature proved to be an interactive conversation partner who helped students develop both their speaking and listening skills by facilitating real conversational exchanges. Through its feedback, Doubao also encouraged vocabulary expansion, allowing users to try new words and expressions in contexts and making language learning feel more natural and less like rote memorization.

During speaking assessment practices, Doubao played a significant role in boosting students' speaking capabilities. The platform served as a friendly interlocutor, making learners more willing to communicate and less afraid of making mistakes. This supportive environment was described by participating students as helpful, interesting, and comforting, working as a patient, judgement-free partner to talk with, helping to overcome anxiety and fear often experienced in real-life language exchanges. The participating students also revealed in their reflective audios that the app helped them with their practice by providing answers to questions and even suggesting appropriate vocabulary and discussion prompts. The discussion below illustrates in more detail how participating students used Doubao for

developing their oral skills. It also shows how Doubao supported them to boost their tutorial discussion and spoken reflection skills.

### ***Findings from qualitative data: Doubao enhances oral skills development***

Using Doubao enhances students' oral skills in the following ways:

#### *By modelling.*

Doubao effectively modelled language expressions by providing an interactive, generative, and adaptive environment where learners could practise pronunciation, expand vocabulary, and refine expressions through real-time interaction and feedback. Doubao simulated discussion scenarios encouraged them to use complex vocabulary and academic phrases, gradually enhancing their ability to express nuanced ideas. Students also noted that repeated practice with Doubao improved their fluency and reinforced confidence, as the model provided by Doubao acted as a judgement-free conversational partner. By introducing new words and modelling natural sentence structures, Doubao facilitated both lexical development and pragmatic use of English in context. One of the students reflected that

Doubao gives more explanations which improves my fluency. In terms of vocabulary, responses sometimes include words and unfamiliar ways of saying. By looking them up, I can gradually build my vocabulary. It also provides a targeted answer to my question, which allows me to practise specific situations. (WR, class 1, week7)

This finding is in line with Singh and Namin's (2025), and Lee et al.'s (2025) who concluded that simulated use of chatbots played an important role in improving students' speech capabilities. In this way, the process of modelling through practice allowed learners to learn authentic language patterns, strengthening their overall oral competence.

#### *By boosting students' confidence.*

This study also found that practicing with Doubao significantly boosted confidence, a point echoed in responses in many reflection audios (WR 8, WR 3, WR 16, WR 20, and WR 22), with the majority of students describing Doubao as an approachable friend that made them less afraid of making mistakes in speech. A student remarked in their reflection that "It has helped me boost my confidence, expand my vocabulary, and improve my ability to express myself. It may not sound as natural as human, but it still plays a positive role in my learning progress" (WR, class 1 week 7). An increasing level of confidence was also reported by the participating students. For example, a student from class 2 reported in week 7 that the app "helps build confidence in speaking by providing instant response/feedback" (WR, class 2, Week 7), and "it helped in boosting my confidence by providing discussion tips, vocabulary

and techniques” (WR, class 1, week 6), and “it helped raise my confidence level by expanding my vocabulary and improving my ability to express myself” (WR, class 2, Week 8).

The platform’s empathetic and judgement-free exchanges, as documented in the audio reflections, helped students overcome anxiety and foster a willingness to communicate, thus making practice sessions both helpful and interesting. This finding support similar findings in the previous studies by Liu et al. (2025), and Du and Daniel (2024). Additionally, the students’ reflection is consistent with the assessor’s notes on their confidence as the table below shows.

Table 1. Assessors’ notes on students’ performances: Confidence improvement

Diagnostic Speaking (week 1)	Tutorial Discussion (Week 6)	Spoken Reflection (week 9)
10 out of 14 (5 + 5 =2 classes)	14 students from 2 classes	14 students from 2 classes
Assessors’ notes suggested at least 5 experimental students from each class had “a nervous start” or “hesitated a lot” or “lacked confidence”.	Assessors noted “confident” in most cases	Assessors noted “confident to very confident” in most cases

Table 1 above illustrates that the assessors noted an increased level of confidence in students as the use of Duobao progressed through the course. The reasons students noted for increased confidence include “overcoming their fear of speaking”, and “having someone friendly to talk to without being judged”. The assessors also noted that the five concerned students were speaking more and were able to sustain their discussions.

*By improving speech quality.*

When it comes to improving speech quality - an important assessment criterion for marking/grading ME 4 tutorial discussion and spoken reflection - Doubao was praised by participating students for fostering fluency and providing crucial pronunciation feedback. They mentioned being able to polish their spoken English and remedy mistakes as the app actively corrected grammatical and pronunciation errors. In addition, similar to findings in Shikun et al. (2024), it helped learners develop their perspectives through conversations and feedback, encouraging them to think beyond merely repeating text answers and encouraging more in-depth, critical engagement with topics/theory.

The platform played a clear role in improving speech quality, as detailed in their weekly reflections (R3, R6, R21, and R14). Students reported notable improvement in fluency, pronunciation, and the ability to structure ideas, thanks to instant feedback and repeated correction. Weekly reflection audio transcripts highlighted learners’ progression in

expressing their own perspectives and deepening their understanding through Doubao's diverse conversational prompts. Thus, Doubao provided substantial support for developing oral skills, as reflected in user feedback from focus group interview and many Reflective File Sources particularly reflections from R1, R10, R8, R9, and R4, by enabling frequent practice and naturally stimulating users to speak in a more engaging way. One student, for example, reflected that

I had many conversations about the theory of hidden needs, Packard's theory and self-directed learning. And I think it helps. Because when I talk, I can talk whatever I want, it gives an instant response. And mostly it works just fine, so I think it's really efficient for me to improve my speaking skills (WR, class 2, week 1).

Another student reflected that "it helps me find alternative ways to explain ideas when I cannot directly express them in English" (WR, class 1, week 7). Similar comments came from other students' reflections as follows:

I asked Doubao to simulate discussion situations with me on the topic of 5 hidden needs. I described some advertisements to it and talked about their purpose and hidden needs used. Doubao asked me some details and showed its own perspectives. I think Doubao is not only a generative AI but also can be a good friend to chat with and a nice teacher to learn from. Talking to it is like having a tireless judgement free partner who answers all your questions and provokes your new thoughts. It put forward many insightful viewpoints that will greatly inspire me in the future discussion. All in all, I think Doubao is really helpful (WR, class 2, week 5).

*By providing instant feedback.*

Participants further praised Doubao as a valuable aid and a practical, and efficient tool. The instant feedback and practical conversational exercises made the tool highly convenient for fragmented, self-directed language study (Shikun et al. (2024). A student (from class 1) provided their reflection in this way:

This is my audio reflection for week 7. I've been practising with Duobao for 6 weeks and I think this practise did help me a lot. Firstly, it helps me a lot with my vocabulary during my speech, particularly my tutorial discussion practice. I mean I can just ask for whatever I want. For example, I can ask what I am going to use during your speech. Moreover, it helps me a lot with my understanding when it comes to the new theory, like what acculturation and deculturation and tells me how to use them. It also tells me what examples I am going to use and so on.

Students reported that Doubao's instant corrections on grammar and pronunciation helped them identify and fix mistakes, leading to clearer and more accurate speech production. As a reliable tool, it worked reliably and could be accessed anytime, making it suited for

fragmented, self-directed learning. Providing instant feedback and practical suggestions, Doubao's design allowed for repetitive practice and immediate correction, making it easy to use consistently for progress in speaking skills.

*By fostering tutorial discussion and spoken reflection (viva voce) skills.*

Doubao helped foster both speaking and listening skills, facilitated genuine conversation, built vocabulary, and made oral practice interesting. Participating students repeatedly mentioned across several reflective audio transcripts that while preparing for their tutorial discussion and spoken reflection assessments, they felt more comfortable speaking and experimenting with spoken English in simulated conversational settings. Many students reported that the app helped them in several ways. A participant reflected:

Initially, I created a scenario, a simulated scenario where I can be the one who presents their presentations. So, when I'm not sure about some questions, I ask questions to let AI give me answers. On most occasions, AI works very well. But for some of the questions I have to present lots of settings (prompts), like the background of my topic of the presentations. So, once I got the answers from the AI, I am ready to present (WR, class 1, week 1).

Another student noted that "it helped me continually polish my speaking skills, and my answers for the preparation of the presentations" (WR, class 1, week 7).

*By preparing students for speaking assessment.*

The students found the Doubao app useful for assessment purposes as well. The use of Doubao enabled students to evaluate their own speaking abilities and offered judgment on pronunciation and correctness of answers, often functioning as both a rehearsal partner and an evaluator before real discussions or exams. This kind of targeted feedback prepared them for academic conversations and helped them identify areas for improvement.

Regarding their own assessment, a number of students (WR 21, WR 24), and the focus group interview stressed that Doubao helped users effectively evaluate their own speaking ability, highlighting pronunciation errors, checking answer correctness, and preparing students for tutorial discussions and oral exams. Students' reflections also cited that it gave targeted, actionable feedback pre-discussion, as frequently cited in audio and interview sources.

In general, qualitative data findings shows that students were deeply engaged by Doubao's capacity to model tutorial scenarios and simulate question and answer exchanges, which significantly enhanced their mastery of the targeted skills. Beyond offering practice opportunities, the chatbot provided valuable insights into the structures and techniques of both discussion and reflection formats. In hindsight, researchers concluded that Doubao had

evolved from a mere learning aid into a collaborative partner, one that could act as an objective, non-judgemental assessor, giving students constructive feedback on their performance while reinforcing their confidence and competence.

### ***Findings from test results***

Findings from qualitative data were supported by findings from quantitative data from the test results. Table 2 illustrates changes in speaking performance between the experimental and control groups across three assessment points. Specifically, it compares Class 2 students' results on the diagnostic speaking (week 1), tutorial discussion (week 6), and spoken reflection (week 9). In the diagnostic speaking stage, the control group scored slightly higher (72%) than the experimental group (69%), indicating a small initial advantage in oral proficiency. However, by the tutorial discussion and spoken reflection stages, both groups achieved identical scores (80%), suggesting that the experimental group had caught up over time. This parity in later stages implies that the experimental intervention, possibly involving additional support or use of a specific learning tool, was effective in improving the experimental group's speaking performance to the level of the control group.

Table 2. Class 2 comparison of average results from the diagnostic speaking (week 1), tutorial discussion (week 6) and spoken reflection (week 9)

	Experimental Group	Control Group	Remarks
Diagnostic Speaking (Week 1)	69%	72%	Control group score is higher by 3 points
Tutorial Discussion (Week 6)	80%	80%	Experimental group and Control group's scores are at par.
Spoken Reflection (presentation + Q&A) (Week 9)	80 %	80%	Experimental group and Control group's scores are at par.

Table 3 presents a comparison of class 1's speaking performance between the experimental and control groups across three assessment stages: diagnostic speaking (week 1), tutorial discussion (week 6), and spoken reflection (week 9). At the initial stage, both groups displayed similar proficiency levels, with the experimental group scoring slightly higher (53%) than the control group (52%). By week 6, the experimental group's score increased notably to 80%, compared to the control group's 78%, showing clear progress in oral communication skills. In the final stage, spoken reflection, the experimental group achieved a substantial improvement, reaching 91%, while the control group scored 86%. This consistent upward trend indicates that the experimental intervention had a strong positive impact on

students' speaking abilities, particularly in fluency, confidence, and reflective communication performance.

Table 3. Class 1 comparison of average results from the diagnostic speaking (week 1), tutorial discussion (week 6) and spoken reflection (week 9)

	Experimental Group	Control Group	Remarks
Diagnostic Speaking (Week 1)	53%	52%	The scores suggest that the experimental group constantly made significant improvement.
Tutorial Discussion (Week 6)	80%	78%	
Spoken Reflection (presentation + Q&A) (Week 9)	91%	86%	

Table 4 compares the speaking performance of the experimental and control groups between the diagnostic speaking test and the combined EDP speaking test components (tutorial discussion and spoken reflection). Initially, both groups began with similar proficiency levels, as shown by close diagnostic speaking scores (experimental 64%, control 62%). By the tutorial discussion stage, both groups improved noticeably, with the experimental group slightly ahead (80% vs. 79%). In the final spoken reflection stage, the experimental group maintained its lead, achieving 86% compared to the control group's 83%. This progression demonstrates that while both groups improved over time, the experimental group showed a more pronounced gain, suggesting that the experimental intervention effectively enhanced learners' overall speaking proficiency, confidence, and ability to articulate ideas in extended spoken tasks.

Table 4. Diagnostic speaking vs EDP speaking test (TD and SR combined)

	Experimental Group	Control Group
Diagnostic Speaking (Week 1)	64%	62%
Tutorial Discussion (Week 6)	80%	79%
Spoken Reflection (presentation + Q&A) (Week 9)	86%	83%

Table 5 compares the overall/combined speaking performance of the experimental and control groups across two classes. In the diagnostic speaking stage, both groups began at comparable proficiency levels, with the experimental group slightly higher (64%) than the control group (62%). By the speaking tests stage, which includes both tutorial discussion and spoken reflection, the experimental group's score rose to 83%, surpassing the control group's 81%. This improvement indicates that the experimental group achieved greater progress in overall speaking ability, likely due to the effectiveness of the experimental intervention in enhancing fluency, accuracy, and confidence in spoken communication.

Table 5. Overall experimental groups vs control groups speaking performance: Two classes combined

	Experimental Group	Control Group
Diagnostic Speaking	64%	62%
Speaking Tests (TD and SR combined)	83%	81%

### *Challenges when using Doubao*

However, Doubao is not without its limitations. Learning spoken language and discussion techniques extends beyond mastering vocabulary and grammar to include understanding tone, context, and cultural nuance, areas where AI models still fall short (Ma et al., 2025). While AI responses could lack depth and realism, teachers provided adaptive, empathetic guidance that nurtures both linguistic accuracy and essential interpersonal skills like collaboration and critical thinking. Through their human insight and experience, teachers helped learners develop creativity, intercultural awareness, and the ability to engage effectively in real-world, unpredictable conversations.

Participating students pointed out that they sometimes felt that it was just a machine and could not function like a real person. This resulted in emotional detachment and a lack of nuanced human interaction, making it less capable of fully replicating the real conversations. They also noted that it could not solve all speaking problems, especially those involving cultural implications, body language, and emotional subtleties, leading some users to prefer combining Doubao with other learning methods or face-to-face interactions for a balanced approach.

Despite its many benefits, students candidly discussed Doubao's limitations in a number of reflections (WR 6, WR 7, and the focus group interview). Users pointed out that, as a machine, Doubao could not fully replicate real human emotion and nuance, sometimes leading to a sense of disconnection and incomplete feedback. Reflection accounts highlighted that while Doubao solved many practical speaking problems, it could not address all

challenges related to authentic interaction, suggesting its best use as a supplement rather than a sole solution. Some students reflected that:

When it comes to problems, I had during practising with Doubao, I think. I felt like I was talking to a random stranger each time. And I have to repeat everything about a theory again each time. It may be annoying sometimes mechanical, making it obvious that they are not from a real person. (WR, class 1, week 7)

The downside is that sometimes it interrupts me when I pause to think which works my brain to think. Although this is not a serious issue. Actually, I think the fixed problem is that its response lacks the natural tone and style of human conversations (WR, class 1, week 7).

Some other students revealed that “sometimes it does not understand what I say and it can't understand my pronunciation clearly. It frustrates me” (WR, class 1, week 7); or that “though it doesn't fully replicate real human conversation, in general it is a very useful tool for learners, particularly beginners like me” (WR, class 2, week 1).

The researchers observed that students were deeply engaged by Doubao's capacity to model tutorial scenarios and simulate question and answer exchanges, which significantly enhanced their mastery of the targeted skills. Beyond offering practice opportunities, the chatbot provided valuable insights into the structures and techniques of both discussion and reflection formats. In hindsight, researchers concluded that Doubao had evolved from a mere learning aid into a collaborative partner, one that could act as an objective, non-judgemental assessor, giving students constructive feedback on their performance while reinforcing their confidence and competence.

## **Discussion**

The findings from this study reveal that the use of the AI-powered chatbot Doubao has a significant and multifaceted impact on the development of students' English-speaking skills. The main themes that emerged from the data include Doubao's capacity to model authentic language use, boost learner confidence, enhance speech quality, provide instant feedback, and support preparation for assessment tasks, such as discussions and oral presentations. These outcomes were evident in both qualitative reflections and recorded performance data, with students reporting that Doubao offered a supportive, judgement-free environment that encouraged them to practise freely and gain fluency. Unexpectedly, the study also uncovered strong peer collaboration among participants from different classes, with students naturally forming a supportive learning community, which further reinforced their confidence.

Overall, these findings directly address the project's central question concerning how an AI chatbot can enhance oral communication skills in academic English contexts. The results

confirm that Doubao successfully facilitated measurable progress in fluency, accuracy, and communication strategies while reducing anxiety commonly associated with speaking in a second language. These patterns are consistent with existing research by Nguyen (2024), Tai and Chen (2024), and Knoch and Chapelle (2018), which emphasizes the value of AI-assisted learning tools for supporting continuous speaking practice, providing adaptive feedback, and promoting learner autonomy. The present study extends this scholarship by illustrating how AI can be integrated effectively into an EAP learning program, not as a substitute for instruction but as a dynamic complement that bridges gaps in limited classroom time and students' need for individualised feedback.

Despite its strengths, the study acknowledges its certain limitations. Doubao's AI-driven responses, while linguistically accurate, sometimes lacked the depth, emotion, and spontaneity of genuine human conversation, which limited the authenticity of communicative interactions. A small number of participants expressed frustration with the chatbot's mechanical tone and limited emotional nuance, which made interactions feel less authentic. In hindsight, incorporating more teacher-guided reflection sessions alongside chatbot practice could have enhanced students' understanding of cultural and pragmatic dimensions of communication. These findings raise further questions for future research, particularly regarding how AI tools might evolve to capture the subtleties of tone, context, and intercultural communication, as well as the potential long-term effects of consistent chatbot interaction on advanced language learners.

The project has also contributed to reflection on teaching practice, highlighting the potential of AI as a scalable, flexible, and pedagogically valuable resource for improving speaking proficiency. It underscored the importance of helping students develop strategies to use generative AI critically and purposefully, ensuring that they view such tools as complements to, rather than replacements for human interaction. The experience reinforced the idea that blended approaches, where AI-based support is layered with human feedback, offer the most promising pathway to developing both linguistic competence and communicative confidence in EAP settings.

## **Conclusion**

The findings from this study have important implications for enhancing speaking pedagogy in English for EAP and ELICOS programs within our institute (and beyond). They demonstrate that AI-powered tools such as Doubao can effectively complement classroom instruction by providing frequent, low-pressure opportunities for oral practice, personalized feedback, and confidence building. The greatest impact is likely to be seen in contexts where students have limited access to one-on-one speaking support or experience high levels of speaking anxiety. At our institute, these insights could inform the integration of structured AI-speaking practice into existing curricula, particularly within our English programs. Beyond our organization, other institutions could adopt similar models to promote learner

autonomy and ensure equitable access to language practice outside class time. Moving forward, the next stage of this action research will involve refining the approach by exploring collaborative AI-human feedback phases, developing classroom frameworks for guided use of generative AI, and evaluating the long-term transfer of AI-supported speaking gains to authentic academic and professional communication contexts.

## **Acknowledgements**

We acknowledge the contribution of the ME 4 students in Melbourne and the EDP students in Suzhou for their participation with the research. We also acknowledge the contribution of the local teacher(s) in the EDP programme in Suzhou.

We acknowledge the use of Gen AI sparingly to improve the readability of text.

## References

- Al-Husban, N. A. (2025). The impact of AI-assisted language learning tools on augmenting university EFL students' speaking skills in Jordan. *Journal of Applied Learning & Teaching*, 8(1), 116–127. <https://doi.org/10.37074/jalt.2025.8.1.13>
- Alwin, W. R., Manurung, K., Lebagi, D., & Mukrim, M. (2024). The use of English Language Speech Assistant (ELSA) speak application to improve pronunciation. *Jurnal Onoma: Pendidikan, Bahasa, dan Sastra*, 10(3), 2880–2887. <https://doi.org/10.30605/onoma.v10i3.3926>
- Anguera, X., Proença, J., Gulordava, K., Tarján, B., Parslow, N., Dobrovolskii, V., & Girard, R. (2023). ELSA speech analyser: English communication assessment of spontaneous speech. In *Proceedings of the 9th Workshop on Speech and Language Technology in Education (SLaTE)* (pp. 95–96). Association for Computational Linguistics.
- Chae, J. (2026, December 31). *How much does a language tutor expert cost?* Expert Sapiens. <https://expertsapiens.com/en/cost/language-tutor/>
- Chen, Y.-L., Hsu, C.-C., Lin, C.-Y., & Hsu, H.-H. (2022). Robot-assisted language learning: Integrating artificial intelligence and virtual reality into English tour guide practice. *Education Sciences*, 12(7), Article 437. <https://doi.org/10.3390/educsci12070437>
- 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, Article 100190. <https://doi.org/10.1016/j.caeai.2024.100190>
- El Shazly, R. (2021). Effects of artificial intelligence on English speaking anxiety and speaking performance: A case study. *Expert Systems*, 38(3), e12667. <https://doi.org/10.1111/exsy.12667>
- Han, D. E. (2020). The effects of voice-based AI chatbots on Korean EFL middle school students' speaking competence and affective domains. *Asia-pacific Journal of Convergent Research Interchange*, 6(7), 71-80.
- Hoang, N. T., Han, D. N., & Le, D. H. (2023). Exploring chatbot AI in improving vocational students' English pronunciation. *AsiaCALL Online Journal*, 14(2), 140–155. <https://doi.org/10.54855/acoj.231429>
- Jeon, J., Lee, S., & Choe, H. (2023). Beyond ChatGPT: A conceptual framework and systematic review of speech-recognition chatbots for language learning. *Computers & Education*, 206, 1-19. <https://doi.org/10.1016/j.compedu.2023.104898>
- Karnieli-Miller, O., Strier, R., & Pessach, L. (2009). Power relations in qualitative research. *Qualitative Health Research*, 19(2), 279–289. <https://doi.org/10.1177/1049732308329306>
- Lee, Y. J., Proctor, M., Roger, P., Enriquez Raido, V., & Dras, M. (2025). *Designing data-responsive language learning environments through conversation practice with AI* (Research project report). Department of Linguistics & School of Computing, Macquarie University. <https://researchers.mq.edu.au/en/projects/designing-data->

[responsive-language-learning-environments-through-/](#)

- Liu, J., Hamid, H. A., & Bao, X. (2025). Motivation and achievement in EFL: The power of instructional approach. *Frontiers in Education*, 10, 1-12. <https://doi.org/10.3389/feduc.2025.1614388>
- Ma, M., Noordin, N., & Razali, A. B. (2024). Effects of an AI chatbot mobile application on foreign language anxiety among Chinese EFL undergraduates. *International Journal of Academic Research in Progressive Education and Development*, 13(4), 3828-3840. <https://doi.org/10.6007/IJARPED/v13-i4/23894>
- Makhlouf, M. K. I. (2021). Effect of artificial intelligence-based application on Saudi preparatory-year students' EFL speaking skills at Albaha University. *International Journal of English Language Education*, 9(2), 36-57 <http://ijele.macrothink.org>
- Mohamed, A. M. (2024). Exploring the potential of an AI-based chatbot (ChatGPT) in enhancing English as a foreign language (EFL) teaching: Perceptions of EFL faculty members. *Education and Information Technologies*, 29, 3195-3217. <https://doi.org/10.1007/s10639-023-11917-z>
- Mohammadkarimi, E. (2024). Exploring the use of artificial intelligence in promoting English language pronunciation skills. *LLT Journal: A Journal on Language and Language Teaching*, 27(1), 98-115. <https://doi.org/10.24071/llt.v27i1.8151>
- Nguyen, H. A. (2024). Harnessing AI-based tools for enhancing English speaking proficiency: Impacts, challenges, and long-term engagement. *International Journal of AI in Language Education*, 1(2), 18-29. <https://doi.org/10.54855/ijaile.24122>
- Nguyen, H., H. (2025). AI for English speaking practice: A study of effectiveness and engagement among Vietnamese university learners. *Australasian Journal of Educational Technology*, 41(3), 45-61. <https://doi.org/10.14742/ajet.10050>
- Pokrivčáková, 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>
- Qiao, H., & Zhao, A. (2023). Artificial intelligence-based language learning: Illuminating the impact on speaking skills and self-regulation in the Chinese EFL context. *Frontiers in Psychology*, 14, 1-15. <https://doi.org/10.3389/fpsyg.2023.1255594>
- Shikun, S., Grigoryan, G., Ning, H., & Harutyunyan, H. (2024). AI chatbots: Developing English language proficiency in the EFL classroom. *Arab World English Journal (Special Issue on ChatGPT)*, 292-305. <https://doi.org/10.24093/awej/ChatGPT.20>
- Singh, S. U. & Namin, A. S. (2025). A survey on chatbots and large language models: Testing and evaluation techniques. *Natural Language Processing Journal*, 10, 1-18. <https://doi.org/10.1016/j.nlpj.2025.100004>
- Tai, T. Y., & Chen, H. H. J. (2024). Improving elementary EFL speaking skills with generative AI chatbots: Exploring individual and paired interactions. *Computers & Education*, 220, 1-5. <https://doi.org/10.1016/j.compedu.2024.105112>
- Thomas, C. G. (2021). *Research Methodology and Scientific Writing* (2nd ed.). Springer.
- Yusuf, H., Money, A. & Daylamani-Zad, D. (2025). Towards reducing teacher burden in

performance-based assessments using aivalue: An emotionally intelligent LLM-augmented pedagogical AI conversational agent. *Education and Information Technology*, 30, 24649-24693. <https://doi.org/10.1007/s10639-025-13755-7>

Zhu, Y., Zou, B., & Jia, W. (2025). The impact of artificial intelligence (AI) technology on EAP speaking. *International Journal of Distance Education Technologies*, 23(1), 1-19. <https://doi.org/10.4018/IJDET.380730>

Zou, B., Liviero, S., Hao, M., & Wei, C. (2020). Artificial intelligence technology for EAP speaking skills: Student perceptions of opportunities and challenges. In M. R. Freiermuth & N. Zarrinabadi (Eds.), *Technology and the psychology of second language learners and users: New language learning and teaching environments* (pp. 433–463). Palgrave Macmillan. [https://doi.org/10.1007/978-3-030-34212-8\\_17](https://doi.org/10.1007/978-3-030-34212-8_17)

---

**Saman Ranabahu** is an EAP teacher at Monash College with over 25 years of experience in a range of roles, including coordinator, independent learning specialist, leading teacher, assessor, trainer, and EAP/ELICOS educator. He holds a Master's degree in Education from the University of Exeter, UK. His current research interests focus on learner development, the use of AI in language learning, and formative assessment.

**Dr Ram Ashish Giri** is an academic at Monash College where he has been teaching in the English as well as Diploma programs since 2013. Dr Giri received his postgraduate education from Columbia University, New York and PhD from Monash University. He has researched and published on a wide range of English language education issues. His most recent co-edited books include (a) *English as a Medium of Instruction in South Asia: Issues in equity, equality and social justice* (Routledge, 2024) (b) *English as a Medium of Instruction: Case studies from Asia* (Springer, 2025) and (c) *The Handbook of English Language Education in Nepal* (Routledge, 2025).

**Dr Alastair James** is a Program Leader of Monash English, having worked at Monash College for 13 years.