



## Comparative Evaluation of Peer-Led and AI Mediated ESL Speaking Skill Development: Insights from Sri Lankan University Learners

P.B.S.L. Pushpakumara<sup>1</sup>, T.A.S. Sudehi<sup>2</sup>, K.S.P Dias<sup>3</sup>

### Abstract

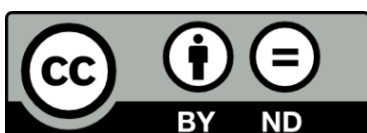
This study followed a mixed-methods approach to explore how Sri Lankan undergraduates viewed the use of ChatGPT Voice Mode (VM) and peer interaction (PI) in improving English-speaking proficiency. After experiencing both methods for one year, a questionnaire was distributed to 42 English as a second language (ESL) students to assess the limitations and strengths of each method. Preferences were almost evenly split when the participants had to choose between the two. ChatGPT VM was preferred for its non-judgemental and safe atmosphere, which helped students develop accuracy, confidence and reduce speaking anxiety. Furthermore, the participants stated that ChatGPT VM supported structured learning, particularly in grammar and vocabulary. As limitations of ChatGPT VM, complex answers provided, lack of naturalness and technical issues were cited. Difficulty in comprehending pronunciation was viewed as another limitation. PI was preferred for its social quality and its ability to provide fluent, confident, and real-life conversations. Emotional support, authenticity and practical feedback were also cited as benefits of PI. However, minor concerns were raised regarding the varied proficiency levels of interlocutors. Therefore, as practical suggestions, educators and curriculum developers can integrate each method as complementary tools rather than mutually exclusive alternatives in order to enhance students' overall learning outcomes.

**Keywords:** AI Mediated ESL Speaking, ChatGPT VM, Peer-Led Interaction, Sri Lankan University ESL Learners

<sup>1,2,3</sup>Department of English Language Teaching and International Center for Multidisciplinary Studies, University of Sri Jayewardenepura, Nugegoda, Sri Lanka.

[sampath@sjp.ac.lk](mailto:sampath@sjp.ac.lk)

 <https://orcid.org/0000-0003-0773-2805>



This article is published under the Creative Commons CC-BY-ND License (<https://creativecommons.org/licenses/by-nd/4.0/>). This license permits to use, distribute, and reproduce the contents of the publication for commercial and non-commercial purposes, provided that the original work is properly cited and is not changed anyway.

**Original Article****INTRODUCTION**

English as a second language (ESL) learners believe speaking is one of the key parts of communicative competence. It is also considered the most difficult skill for them to develop, especially in contexts where there are limited opportunities for real-life speaking. Strong speaking skills have become essential for non-native English learners since English has become globally accepted and functions in education (Celik et al., 2025). PI, which is a traditional method of teaching speaking, helps learners improve their fluency, accuracy, and confidence in speaking through negotiation of meaning and feedback (Swain, 2000; Mackey & Gass, 2014). However, whether PI is successful or not depends on the tasks that are used in the classroom, learner needs and institutional limitations (Grab, 2025).

In this technologically driven world, conversational AI tools such as ChatGPT Voice Mode (VM) are increasingly being used in the field of English Language Teaching (ELT) to create various opportunities to improve learners' speaking proficiency. These tools usually offer low-anxiety, personalised support with feedback, and their use of varied prompts to generate content also develops learners' motivation, fluency, grammar, vocabulary, and pronunciation (Nuñez et al., 2025). However, there are clear limitations (Shalihah, 2025), as most studies have only focused on the short-term effects

of technology in English as a Foreign Language (EFL) setting.

This research gap is particularly relevant to the Sri Lankan context, where despite years of formal English education, learner proficiency remains low (Education First, 2025). Undergraduates frequently report struggling with speaking anxiety and a lack of opportunities for real-life interactive communication (Chamba & Hansson, 2024), which are the precise challenges that conversational AI tools claim to address. Thus, Chamba and Hansson (2024) propose that both PI and digital tools provide low-pressure speaking opportunities. However, while some local research has explored peer learning (Pushpakumara, 2024), studies that particularly investigate its influence on speaking development remain limited, and empirical research on AI-mediated speaking practice involving Sri Lankan undergraduates is scarce.

Although previous studies have explored PI and AI-assisted speaking practice, these approaches have largely been investigated separately. Moreover, they have been concentrated mostly within EFL contexts involving short-term interventions. Comparative studies examining both methods within the same learning context remain limited, and little is known about how Sri Lankan undergraduates perceive their relative strengths and limitations for speaking development. This study addresses this gap by comparing Sri Lankan university

**Original Article**

learners' perceptions of PI and ChatGPT VM in speaking development after sustained exposure to both approaches over two academic semesters.

Accordingly, the main objectives of this study are to explore Sri Lankan undergraduates' perceptions of using ChatGPT VM for speaking improvement, investigate their perceptions of PI as a method of speaking practice, and evaluate how they compare the perceived effectiveness of ChatGPT VM and PI as strategies for developing speaking proficiency. Consequently, this study contributes to the growing body of research on AI-assisted language learning by providing evidence from an underrepresented ESL context. Furthermore, it offers practical insights for educators, curriculum designers, and policymakers who are seeking to integrate both collaborative and technological approaches to support speaking development in higher education.

## LITERATURE REVIEW

This literature review examines the importance of speaking in ESL contexts and discusses two approaches commonly used to support speaking development: PI and AI-assisted speaking tools such as ChatGPT VM.

Speaking is considered one of the key communicative competences and essential skills since English is widely

used in the world and globally accepted and used for education, work, travel, and digital communication (Nuñez et al., 2025; Celik et al., 2025). Since millions of people use English globally, effective English-speaking skills allow them to access various opportunities and knowledge. However, learners believe that acquiring speaking skills is difficult due to accuracy, and fluency, highlighting the need for incorporating effective teaching methods in education (Denston, 2021).

ESL teaching uses traditional methods, including guided discussions, presentations, and classroom activities that often rely on PI to improve the speaking skills of learners (Nuñez et al., 2025). Pair work is commonly used in the teaching-learning process where learners can complete tasks like problem-solving and sharing ideas together (Gass & Mackey, 2014). Research done around the world usually highlights that PI supports the development of speaking skills by lowering speaking anxiety and improving confidence and motivation while improving fluency, accuracy, pronunciation, and vocabulary in speaking (Fikry & Zuraida, 2025; Swain, 2000). These benefits are gained by students from authentic communication that involves negotiating meaning and providing feedback. However, the effectiveness of PI still depends on factors such as the way educators design tasks, learner levels and needs, the number of

**Original Article**

students in each class, partner compatibility, and proficiency gaps. Because of that, researchers recommend that studies should focus more on how PI can be used in undergraduate contexts.

Interest in using AI for education grew rapidly with the introduction of ChatGPT in 2022, and with that, educators began to use AI tools, such as chatbots, automated feedback systems, and intelligent tutors with the hope of improving teaching-learning process (Chiu et al., 2023). These technological developments offer various opportunities for language learners, such as helping with speaking practice, which is often limited in traditional classrooms due to time constraints, and helping with learning grammar and vocabulary (Nuñez et al., 2025). AI chatbots are used by educators because they use language-processing tools that allow learners to engage in day-to-day conversations by offering an atmosphere where they feel less pressure in speaking and get consistent and instant feedback for their improvements using tools such as TalkPal AI, Glinglish AI, Langotalk, and ChatGPT VM (Annamalai et al., 2023).

ChatGPT VM recently developed a virtual training environment, built on AI technology using natural language processing tools to assist learners with their conversational practice in a way that provides them with an immersive and low-stress environment (Nuñez, et al. 2025). The main features of ChatGPT

VM that make this possible are (a) simulated dialogue, (b) flexible prompts, (c) instant feedback, (d) language-style control, and (e) autonomous learning (Shalihah, 2025). Most research conducted on ChatGPT VM thus far has focused on EFL learners in Ecuador, Iraq, and Indonesian high school and university learners, typically using a predominantly quantitative or mixed-methods approach. Studies that provide evidence that ChatGPT VM has improved pronunciation, fluency, vocabulary, grammar, and speaking confidence, decreased anxiety, and increased student engagement all report similar results (Celik, et al. 2025; Nuñez, et al., 2025; Salsabil, et al. 2025). In addition to the previously stated advantages of ChatGPT VM, Pratiwi et al. (2024) also mentioned the importance of access to personalised feedback and diverse conversational topics. Although very limited, emerging ESL-related studies (Grab, 2025) also indicate similar advantages.

Limitations of AI-based speaking tools are also reported by previous research. For instance, Grab's (2025) research indicates that there are limitations to using ChatGPT VM to improve speaking, based on data collected from 120 ESL undergraduates from six countries: Brazil, India, Nigeria, Saudi Arabia, Vietnam, and Spain. As mentioned by that study, limitations of using ChatGPT VM are technical issues such as a lack of internet connection and limited technological knowledge;

**Original Article**

teaching-related issues include less authentic practice, time constraints and lack of knowledge of teachers about using technology in the learning and teaching process, and ethical concerns (Shalihah, 2025; Salsabil et al., 2025). These challenges require researchers to do long-term studies with different types of learners in different countries to analyse the benefits of ChatGPT VM in educational settings to foster authentic conversations (Nuñez et al., 2025; Grab, 2025).

Using peers to interact for speaking development is supported by Vygotsky's Sociocultural Theory, which regards learning as occurring through social interaction within the Zone of Proximal Development (Vygotsky, 1978). The learners work together with the aid of each other to gradually become independent (Fikry & Zuraida, 2025). At the same time, Long's Interaction Hypothesis states at the same time, that meaning comprehension during communication is one of the factors that enable the learners to change their language and increase their understanding (Long, 1983, 1996; Gass & Mackey, 2014). This is also true for AI tools, since ChatGPT VM does offer feedback that is immediate, as well as opportunities for negotiating (Grab, 2025). Swain's Output Hypothesis backs up both approaches, as producing language has been pointed out as the main reason why learners can notice the gaps, test their ideas, and refine their skills, which ChatGPT VM does through

meaningful output (Swain, 1995, 2005; Shalihah, 2025).

Overall, while previous studies have demonstrated the benefits of both PI and AI-assisted speaking practice, comparative studies investigating both approaches within a single ESL context remain limited. Therefore, this study seeks to address that gap by exploring how Sri Lankan undergraduates perceive and compare PI and ChatGPT VM as strategies for developing ESL speaking proficiency.

## RESEARCH METHODOLOGY

### Research Design

This study follows a mixed-methods approach, gathering and combining both quantitative and qualitative data that were collected from a questionnaire exploring views of undergraduates on using PI and ChatGPT VM in order to improve their English-speaking skills. Consequently, this study analyses both measurable factors such as comfort, anxiety, motivation, feedback, and comprehension and students' personal opinions and experiences about both PI and ChatGPT VM. Likert-scale questions were used to analyse quantitative data, while qualitative data were gathered through open-ended questions asking students to descriptively write down their experiences regarding the use of ChatGPT VM and PI. The mixed-methods approach was selected to

**Original Article**

obtain a comprehensive understanding of students' perceptions by combining numerical trends with detailed personal experiences.

### Participants and Context

Forty-two (n=42) first-year undergraduates from the Faculty of Humanities and Social Sciences at a state university in Sri Lanka participated in this study. This study employed a simple random sampling technique to ensure greater representativeness and to enhance the generalizability of the findings. Accordingly, each individual in the population had an equal and independent chance of being selected.

The total population comprised 1,373 registered students, from which a sample of 42 students (n=42) was drawn using a computer-generated random number procedure. This sample represents approximately 3.06% of the total population. The adoption of simple random sampling helped to minimise selection bias and strengthen the external validity of the study, thereby making the findings more reflective of the broader student population.

A placement test conducted at the beginning of their academic year indicated that most of the participants belonged to A1, A2 and B1 proficiency levels. Since all the participants were native Sinhala speakers, they were given the option of answering the open-ended questions in their native

language, with the aim of encouraging detailed and authentic responses. In line with this, the closed-ended questions were also given in both their L1 and L2. Participation was voluntary and all responses were anonymized to ensure confidentiality.

### Data Collection Instruments

Data were collected through a bilingual (Sinhala-English) questionnaire administered digitally via Google Forms. It was structured into three main sections. The first section was used to collect the general demographic and background information, including the students' age, gender, perceived proficiency level of English, and frequency of their English practice with peers. The second section included ten 5-point Likert-scale items, ranging from 1 (Strongly Disagree) to 5 (Strongly Agree). These items were equally divided into two subtopics: oral interaction with a peer and oral interaction with ChatGPT VM. Finally, the third section included four multiple-choice questions and an open-ended question, where they were asked to explain why they preferred their chosen method. They had the liberty to respond to this question either in Sinhala or English. This ensured that language was not a barrier to sharing their genuine opinions. They were also given the option to handwrite their responses on a piece of paper as typing in Sinhala can be challenging. This was done to ensure accessibility and comfort. Subsequently, the Sinhala

**Original Article**

responses were translated into English for analysis.

### Procedure

The study was conducted over two consecutive semesters. During the first semester, students used ChatGPT VM to practise their English-speaking skills both during classroom activities and independently outside class. Learners engaged in twelve speaking activities over the semester, with each session lasting around 20-30 minutes. They interacted with ChatGPT VM on a range of familiar topics, such as daily routines, hobbies, personal experiences, university life, and study-related issues, and received real-time feedback and responses from ChatGPT VM. During the second semester, the participants were involved in PI tasks integrated into their regular English lessons. These were also conducted once a week and lasted around 20-30 minutes per session. Accordingly, the participants took part in ten peer-interaction sessions. Here, students engaged in pair or small-group discussions on topics similar to those they used in semester one to ensure comparability between the two phases. The course instructor facilitated all the activities, provided students with guidance where necessary, and monitored their participation.

After the exposure to both approaches, students were invited to reflect on and compare their experiences. The questionnaire was administered at the end of the second semester. After

completing the questionnaire during the class, students were asked to hand over their written responses, and the researchers collected them separately. Before collecting data, participants were informed about the purpose of the study and was assured that their responses would remain confidential and would be used only for research purpose.

### Reliability and Validity of the Instrument

The questionnaire was developed based on the objectives of this study and insights gained from the literature review. The items of the questionnaire were designed to measure key constructs relevant to the study, including speaking anxiety, comfort, motivation, confidence, feedback, comprehension, and learner preferences. To further establish content validity, the questionnaire was evaluated by an experienced academic in the field of ELT. Subsequently, revisions were made to improve the clarity, relevance, and appropriateness of the items based on the feedback received. Moreover, the questionnaire was prepared in both English and Sinhala to minimize potential misunderstandings and ensure that participants could respond accurately. Furthermore, the inclusion of both closed-ended and open-ended questions allowed for triangulation, which further improved the validity of the findings. In terms of improving reliability, the study used a mixed-



## Original Article

methods approach by collecting both quantitative and qualitative data, as suggested by Creswell and Poth (2023). To enhance the reliability of the qualitative findings, two researchers independently coded the responses and resolved any discrepancies through discussion and consensus. Overall, these measures together helped improve the overall reliability and validity of the instrument.

### Data Analysis

Quantitative data obtained from the Likert-scale and multiple-choice questionnaire items were analysed using descriptive statistics such as percentages, means and standard deviations. The findings were then presented through tables to facilitate easy interpretation and comparison of responses. On the other hand, qualitative data collected through open-ended questionnaire items were analysed thematically following the procedure proposed by Braun and Clarke (2021). Subsequently, the analysis involved familiarization with the data, coding of responses, identification of recurring patterns and generation of themes related to learners' experiences with both approaches.

### Ethical Considerations

Voluntary participation was maintained, and students were assured that their responses would not affect their grades. Identifiable and personal information was removed from all

datasets. Participants were also asked to respond either in Sinhala or English to ensure they felt comfortable during the data collection process.

## RESULTS AND DISCUSSION

This section presents the findings from the questionnaire to address the three research questions of the study.

Around 85.7% of the participants agreed that they sometimes practised English with their peers, highlighting that PI is familiar even though it is not done regularly. Around 7.1% said that they engaged in practising English with their peers either never or often, showing that the majority of the participants sometimes utilise PI in their daily lives to improve their English language proficiency.

Students' familiarity with ChatGPT VM was more evenly distributed. 52.4% of students had used ChatGPT VM before whereas, whereas 47.6% had never tried it. This highlights that AI-based tools are new in Sri Lankan undergraduate ESL context. Therefore, this study was able to capture perspectives from both experienced and novice students to provide a deeper understanding of how AI-assisted speaking tools are used among learners.

### Perceptions of ChatGPT Voice Mode

The first research objective was to explore Sri Lankan undergraduates' perceptions of using ChatGPT VM to



## Original Article

improve their English-speaking proficiency (Table 1). Overall, the findings revealed that undergraduate ESL learners generally perceived ChatGPT VM as a supportive and useful tool for developing their speaking skills. However, at the same time, they have identified several limitations related to it as well.

Table 1. Students' Perceptions of using ChatGPT VM to develop Speaking Proficiency

| Perception Item                                                | Mean | Standard Deviation (SD) | Interpretation                      |
|----------------------------------------------------------------|------|-------------------------|-------------------------------------|
| 1. I feel comfortable speaking English with ChatGPT.           | 3.67 | 0.95                    | Agree / Moderate-High Comfort       |
| 2. I receive useful feedback from ChatGPT.                     | 3.79 | 0.78                    | Agree / Value AI Feedback           |
| 3. I feel less anxious speaking to ChatGPT than to a human.    | 3.40 | 1.08                    | Neutral-Agree / Reduced Anxiety     |
| 4. I can clearly understand ChatGPT's voice and pronunciation. | 3.00 | 1.15                    | Neutral / Mixed Comprehension       |
| 5. I feel motivated to practice more after speaking            | 3.50 | 0.89                    | Neutral-Agree / Moderate Motivation |

with  
ChatGPT.

Source: Questionnaire Survey, 2025

The quantitative findings indicated generally positive perceptions of ChatGPT VM. Students stated that they felt comfortable speaking English with ChatGPT ( $M = 3.67$ ,  $SD = 0.95$ ) and agreed that it provided them with useful feedback ( $M = 3.79$ ,  $SD = 0.78$ ). Moreover, they reported that speaking with ChatGPT reduced their anxiety to some extent ( $M = 3.40$ ,  $SD = 1.08$ ) and motivated them to continue practicing speaking in English ( $M = 3.50$ ,  $SD = 0.89$ ). However, responses regarding the comprehensibility of ChatGPT's voice and pronunciation were more mixed ( $M = 3.00$ ,  $SD = 1.15$ ), which suggested that they experienced both advantages and challenges when interacting with ChatGPT VM.

Several interrelated themes that explain these perceptions were revealed through qualitative analysis. One of the major themes among them was ChatGPT as a non-judgemental, and safe learning space. Many students appreciated the opportunity to speak without the pressure of being humiliated by others or the fear of making mistakes when talking in real time. Responses such as "I can talk freely without getting shy" and "Talking with ChatGPT was a wonderful activity because we can make errors without fear" demonstrate how learners valued the absence of judgement when engaging in spoken communication. This aligns with the

**Original Article**

findings of Nuñez et al. (2025) and Salsabil et al. (2025), who claimed that the use of AI tools such as ChatGPT VM created low-tension, personalised support with feedback. Such environments, as Krashen and Terrell (1983) argued in their affective filter hypothesis, ensure better comprehension and language acquisition as learners tend to focus more on meaning-making rather than error-making thanks to reduced anxiety. In the Sri Lankan context, where learners generally experience fear of negative evaluation, failing, public embarrassment, low self-confidence and negative judgement from teachers or peers (Karunanayake, 2022), this feature appears especially beneficial.

Another major theme was language development and knowledge expansion. Grammar support was the most frequently recognised benefit among undergraduates. For instance, a participant stated that "ChatGPT corrects our mistakes when talking. It also uses grammatical sentences." Participants also mentioned that "ChatGPT is best to learn new words" and that they "can improve [their] knowledge day by day," for instance, by using it "to learn meanings of unknown words easily." This suggests that they find ChatGPT useful in improving their vocabulary, pronunciation, and basic grammar knowledge which are essential to form grammatically accurate conversations in their day-to-day lives. It also reveals

how undergraduates perceived ChatGPT as a personalized tutor without solely limiting it to a speaking partner. This exemplifies ChatGPT VM's alignment with Swain's Output Hypothesis (Swain 1995, 2005), which states that producing language is a means for learners to discover their weaknesses in their knowledge about the language. Learners did not only use ChatGPT to practise speaking, but were also able to increase their vocabulary and pronunciation through its use, thus supporting the claims of Pratiwi et al. (2024) that AI provides learners with organised and personalised input into their target language.

Accessibility and immediacy was another recurring theme. Many students appreciated ChatGPT, saying that it is available at any time and can be used anywhere to expand their knowledge. This feature was specially valued by learners who preferred practising at their own pace. One participant stated that "If we respond incorrectly to ChatGPT, we can take ample time to correct our mistakes and speak again." This suggests that ChatGPT VM allowed for controlled practice and self-regulation, which are not always possible in real-time PI. Consequently, it also supports autonomous learning which, according to Nuñez et al. (2025), is a prominent benefit of ChatGPT demonstrated through its inherent features.

Despite these positive perceptions, using ChatGPT VM for speaking practice was not without challenges.



## Original Article

Among them, the most common concern raised was comprehension difficulties. Even though some participants found ChatGPT VM easy to use, others responded that they struggled to understand its pronunciation, vocabulary and speaking speed. For instance, one participant stated that “Its talking speed is too much to follow” while another observed that ChatGPT often “presents difficult words.” These concerns help explain the relatively neutral responses regarding the clarity of ChatGPT’s voice and pronunciation ( $M = 3.00$ ,  $SD = 1.15$ ). It also implies that the linguistic complexity of AI-generated speech may pose challenges for learners of lower proficiency levels.

Technical limitations were identified as another challenge faced by the participants. Some students had difficulty with recognising the AI’s pronunciation of certain words from their own voices. Others emphasised contextually irrelevant responses stating that ChatGPT sometimes had “a very poor understanding of what we say.” More conventional technical problems such as connectivity issues were also reported, with one student noting that “It’s a bit difficult to communicate with ChatGPT when technical mishaps happen.” These findings were corroborated by the results of Grab (2025) and Shalihah (2025) regarding the constraints imposed by the AI’s technical limitations on producing a valid conversational experience.

Overall, students liked ChatGPT VM for its non-judgemental qualities, usefulness, accessibility, and immediacy. They also noticed some challenges, including AI being too complex to be understood by the students and technical issues such as limitations in AI speech recognition and connectivity constraints. Nonetheless, the use of ChatGPT VM to develop speaking proficiency seems to be beneficial in the ESL classroom to reduce pressure and tension and create a conducive language learning environment.

### Perceptions of Peer Interaction

The second research objective of the study was to explore Sri Lankan undergraduates’ perceptions of interacting with peers to improve their English-speaking skills. Overall, students responded positively to using PI, especially in terms of motivation, comfort and the authenticity of communication.

Table 2. Students’ Perceptions of using PI to develop Speaking Proficiency

| Perception Item                                     | Mean | Standard Deviation (SD) | Interpretation              |
|-----------------------------------------------------|------|-------------------------|-----------------------------|
| 1. I feel comfortable speaking English with a peer. | 3.88 | 0.77                    | Agree / High Comfort        |
| 2. I receive useful feedback from my peer.          | 3.86 | 0.75                    | Agree / Value Peer Feedback |



## Original Article

|                                                                  |      |      |                                       |
|------------------------------------------------------------------|------|------|---------------------------------------|
| 3. I feel less anxious speaking to a peer than to an AI Chatbot. | 3.38 | 1.13 | Disagree - Neutral / Moderate Anxiety |
| 4. I can clearly understand my peer's voice and pronunciation.   | 3.79 | 0.92 | Agree / Good Comprehension            |
| 5. I feel motivated to practice more after speaking with a peer. | 3.93 | 0.92 | Agree / High Motivation               |

Source: Questionnaire Survey, 2025

The findings in Table 2 show that most of the students positively reacted to PI when practising English due to various reasons. Accordingly, students reported high levels of comfort when speaking with their peers ( $M = 3.88$ ,  $SD = 0.77$ ) and valued the usefulness of practical and real-time peer feedback ( $M = 3.86$ ,  $SD = 0.75$ ). They also displayed high motivation to continue practising after PI ( $M = 3.93$ ,  $SD = 0.92$ ), even though they were afraid and shy to use their second language in daily conversations. This supports Long's (1996) Interaction Hypothesis, which demonstrates that negotiation of meaning and feedback during interaction facilitate language development.

Further analysis of the qualitative responses reinforced these quantitative findings under three main themes. The first major theme was emotional

warmth and social authenticity. Most of the participants noted the human elements that are present in PI, stating that "Only with humans can we share feelings and emotions" and "Talking with friends is fun and natural." They believed that the presence of humour, expression, and body language plays a critical role in speaking to ensure that they will not be judged or humiliated by anyone when they try to speak in English. Such findings suggest that emotional engagement accelerates learners' willingness to communicate which supports Vygotsky's (1978) argument that social interaction is key to cognitive development.

The second theme was real-life fluency and practical communication. Participants' comments, such as "Talking with friends helps us respond quickly" and "We prepare for real humans one day," demonstrate how they gained an opportunity to improve their fluency through PI. This supports the idea proposed by Swain (2000) that increased interaction through the process of collaborative dialogue promotes increased awareness of linguistic competence and fluency development. Such findings also showed the practicality of natural conversation, particularly how PI trains participants to think and respond spontaneously in English as they engage in regular conversation. Consequently, this unpredictability relies heavily on mutual intelligibility, which is demonstrated by how most participants could clearly understand

**Original Article**

their peers' voice and pronunciation ( $M = 3.79$ ,  $SD = 0.92$ ). In other words, when the barrier of poor comprehension is removed, learners have the mental capacity to practice spontaneous thinking and prepare for real-world interactions.

Finally, the theme of confidence building and fear reduction highlights the gradual development of learners' self-assurance. Although most of the students were initially "very conscious about making mistakes" when talking with their friends, they mentioned that they eventually gained confidence. This is also evident in the quantitative findings, as responses related to anxiety were more varied ( $M = 3.38$ ,  $SD = 1.13$ ). This indicates that while PI promotes confidence, it does not immediately eliminate anxiety. Instead, anxiety fades away with repeated interaction. This reflects Swain's (2000) concept of "pushed output," where learners develop proficiency by being "pushed" to produce language even when it involves some degree of difficulty. Participants also shared that PI allowed them to "reduce [their] fear of public speaking" that they have dealt with for a long time. These findings align with Pushpakumara (2024) and Fikry and Zuraida (2025) who state that that peer learning environments can significantly enhance learner confidence and participation through anxiety reduction.

Although most of the students perceived PI in a positive way, some students pointed out a few minor

challenges. For instance, a few noted that not all peers provided accurate feedback during their speaking sessions. Another concern was that sometimes conversations shifted away from English when trying to understand the meaning of some words. These findings align with the study done by Fikry and Zuraida (2025) who state that the effectiveness of PI depends on a few factors such as partner compatibility, and proficiency gaps.

Overall, PI was considered essential to developing English speaking skills by the participants because it builds fluency and confidence, and offers comfort and authentic speaking practice. Their perceptions reinforce the importance of incorporating speaking practice facilitated by PI in ESL classrooms, particularly in contexts where learners need greater exposure to authentic communication.

### **Comparative Effectiveness of Methods: ChatGPT vs Peer Interaction**

The third research objective of the study was developed to examine how students compare ChatGPT VM and PI in improving their English-speaking proficiency. To directly compare learner preferences, four targeted choice-based items were included in the questionnaire (Table 3). Overall, the findings show that students view both approaches as complementary tools rather than competing alternatives, since both help learners improve their speaking skills in various ways.



## Original Article

Table 3. Distribution of Overall Learner Preferences (n=42)

| Evaluation Criterion            | Preferred PI   | Preferred ChatGPT VM | Preferred Both Equally | Preferred Neither |
|---------------------------------|----------------|----------------------|------------------------|-------------------|
| 1. Overall Speaking Improvement | 15<br>(35.7 %) | 12<br>(28.6 %)       | 14<br>(33.3 %)         | 1<br>(2.4 %)      |
| 2. Anxiety Reduction            | 18<br>(42.9 %) | 14<br>(33.3 %)       | 10<br>(23.8 %)         | 0<br>(0.0 %)      |
| 3. Vocabulary Acquisition       | 8<br>(19.0 %)  | 23<br>(54.8 %)       | 11<br>(26.2 %)         | 0<br>(0.0 %)      |
| 4. Final Forced Learning Choice | 20<br>(47.6 %) | 21<br>(50.0 %)       |                        | 1<br>(2.4 %)      |

Source: Questionnaire Survey, 2025

After analysing students' perceptions both quantitatively and qualitatively, three themes were identified. The first theme was AI for accuracy and peers for fluency. Students often spoke about ChatGPT VM being useful for grammar correction and structured learning, whereas peers were necessary for spontaneous, natural conversation. This is evident in the quantitative

findings as well since, for instance, most learners (54.8%, n=23) preferred ChatGPT VM for vocabulary acquisition while only 19.0% (n= 8) selected PI. This can be attributed to AI's capacity to provide immediate access to varied vocabulary items, precise pronunciation models, and uninterrupted opportunities to practise. Unlike PI which can sometimes be constrained by an interlocutor's proficiency levels, ChatGPT offers a consistently reliable linguistic resource for learners to refer to. In contrast, one participant noted that "We can take as much time as we want to build sentences one by one and talk" when using ChatGPT VM, whereas PI requires more fluency in terms of faster, more spontaneous responses. However, this strong preference for ChatGPT VM for a single function should not be interpreted as a rejection of peer-based learning and vice versa. Instead, it demonstrates learners' strategic thinking which involves efficiency and control when dealing with separate language components.

Moving on, when examining affective factors such as anxiety reduction, a majority of the participants preferred PI (42.9%, n = 18) compared to ChatGPT VM (33.3%, n = 14). This finding complicates the commonly held assumption that AI-mediated environments are inherently less anxiety-inducing. This paved the way for the second theme, which was real humans as the final benchmark. Here,

**Original Article**

even with the value placed on ChatGPT VM by considering it mostly as a stress-free environment, students acknowledged that ultimate communication is human communication. Comments such as “ChatGPT is good, but it's important to talk to humans” indicated that students still value the emotional support and shared experience that comes from PI. This further validates Vygotsky's (1978) Sociocultural Theory which emphasises the importance of human interaction in the learning of language. Comments such as “In our professional lives, we'll have to communicate with real humans and not machines one day” also reflect the findings of Santoso et al. (2025) who pointed out that AI voice tools “...cannot take the place of the subtle, phonetic and pragmatic model” present in human interaction.

Finally, the last theme that was observed is complementary rather than competing. A considerable number of students said that “The best method is to use both together,” which conveyed the sense that ChatGPT VM and PI perform different roles that ultimately support each other. The most noticeable evidence of this complementary perception emerges from the final forced-choice item where learners were asked to choose only one method for their speaking development journey. Here, the cohort split almost evenly, with 50% (n=21) selecting ChatGPT VM and 47.6% (n=20) choosing PI. This indicates that learners demonstrate a clear awareness

of the non-overlapping strengths of each method.

A sub-theme that emerged here was the developmental sequencing of tool use. Several participants commented that they used ChatGPT VM as “a stepping stone” before engaging in PI, and others stated that PI is a useful preparatory tool to practice for real world speaking interactions. This pattern can be interpreted as a form of scaffolded learning where ChatGPT VM provides support for initial confidence building and linguistic accuracy which is then transferred to more socially complex interactions starting with close interlocutors such as peers. This reflects Vygotsky's (1978) concept of the Zone of Proximal Development where learners move from supported output to independent performance through guided practice.

Overall, the comparative findings reveal that learners take a functional and complementary perspective toward ChatGPT VM and PI. ChatGPT VM generally helped them with accurate grammar, vocabulary, and non-judgmental practice, while PI allowed them to practice real-life communication to become more confident and natural when they speak. In other words, ChatGPT VM primarily supports controlled practice whereas PI cultivates communicative competence. This aligns with the fact that AI will never replace the emotional and social aspects of human communication even though it strengthens the accuracy of speaking as found by other

**Original Article**

international researchers (Grab, 2025; Shalihah, 2025). All in all, it is clear that both methods offer a complementary, effective and holistic approach for improving English-speaking proficiency among ESL undergraduates in Sri Lanka and other similar contexts.

Consequently, the present study makes several important contributions to both local and international discussions on AI-assisted speaking development and language learning. Firstly, it is among the few studies conducted in Sri Lanka that directly compare AI-mediated speaking practice and PI in an ESL context. Secondly, it explored learners' perceptions following sustained exposure to both approaches over two consecutive academic semesters, which contrasts with many previous studies that have examined only short-term interventions. Thirdly, the comparative design enabled a more detailed understanding of the strengths and limitations of each approach rather than evaluating each of them in separate studies using a single method. Finally, the findings of this study prove that learners perceive ChatGPT VM and PI as complementary rather than competing methods, and this highlights the potential value of integrating such technological and collaborative learning opportunities to support speaking development in higher education ESL contexts.

**CONCLUSION**

This study examined how Sri Lankan undergraduates perceived ChatGPT VM and PI in developing their English-speaking skills. The findings revealed that students praised ChatGPT VM for lowering their speaking anxiety by offering a safe, non-judgemental atmosphere while supporting grammar and vocabulary. However, they also revealed its limitations, citing complex responses, occasional misunderstandings, pronunciation difficulty, and technical issues. In contrast, students praised PI for being more natural and by allowing them to think for themselves, share ideas and form conversations with real people freely and confidently. Overall, it is clear how students prefer both as complementary tools rather than alternatives. Therefore, both approaches can be used in ESL programmes in universities for undergraduates to improve their speaking proficiency.

The conclusions of this research indicate that ESL teachers should consider employing a combined method involving the use of AI technology and PI techniques to develop students' speaking skills in English language lessons. For practical application, universities can utilise ChatGPT VM in speaking classes as an additional resource to aid r learners in grammar, pronunciation, vocabulary, and stress-free speaking practice. Simultaneously, PI exercises such as pair work, group discussions, role



## Original Article

plays, and other interactive techniques should not be maintained to cultivate their fluency and ability to communicate spontaneously. Future research could examine the long-term effects of incorporating this blended approach into the teaching-learning process and investigate how students use AI tools independently outside the classroom. Additionally, it is useful to investigate how AI tools can enhance the confidence and accuracy of beginner-level ESL learners before developing fluency through peer-led interactions.

## References

- Annamalai, N., Eltahir, M. E., Zyoud, S. H., Soundrarajan, D., Zakarneh, B., & Salhi, N. R. A. (2023). Exploring English language learning via chatbot: A case study from a self determination theory perspective. *Computers and Education: Artificial Intelligence*, 5, 100148. <https://doi.org/10.1016/j.caeai.2023.100148>
- Braun, V., & Clarke, V. (2021). *Thematic analysis: A practical guide*. SAGE Publications. <https://doi.org/10.4135/9781529780494>
- Celik, B., Yildiz, Y., & Kara, S. (2025). Using ChatGPT as a virtual speaking tutor to boost EFL learners' speaking self-efficacy. *Australian Journal of Applied Linguistics*, 8(1), 102418. <https://doi.org/10.29140/ajal.v8n1.102418>
- Chamba, Z. N., & Hansson, H. (2024). Survey: English speaking skills among Sri Lankan undergraduate students: A university case study. In T. Bastiaens (Ed.), *Proceedings of EdMedia + Innovate Learning*. Association for the Advancement of Computing in Education (AACE). <https://www.learntechlib.org/primary/p/226131/>
- Chiu, T. K., Xia, Q., Zhou, X., Chai, C. S., & Cheng, M. (2023). Systematic literature review on opportunities, challenges, and future research recommendations of artificial intelligence in education. *Computers and Education: Artificial Intelligence*, 4, 100118. <https://doi.org/10.1016/j.caeai.2022.100118>
- Creswell, J. W., & Poth, C. N. (2023). *Qualitative inquiry and research design: Choosing among five approaches* (5th ed.). SAGE Publications.
- Denston, A. (2021). Review of the literature on speaking and listening skills: Report to the Ministry of Education (Aotearoa New Zealand). Massey University. <http://hdl.handle.net/10179/17740>
- Education First. (2025). *EF English proficiency index*. EF. <https://www.ef.com/wwen/epi/>
- Fikry, M. A., & Zuraida, N. Z. (2025). Assessing the impact of peer interaction on the English speaking proficiency of the 2nd semester students of English Education Study Program at the Faculty of Teacher Training and Education, Sriwijaya University. *Journal Sains Student Research*, 3(6), 650–662. <https://doi.org/10.61722/jssr.v3i6.6779>
- Gass, S. M., & Mackey, A. (2014). Input, interaction, and output in second language acquisition. In *Theories in second language acquisition* (2nd ed., pp. 194–220). Routledge. <https://doi.org/10.4324/9780203628942-15>
- Grab, M. Ö. (2025). Integrated AI chatbot practice: A pathway to improved ESL speaking skills. *Social Sciences & Humanities Open*, 12, 101933. <https://doi.org/10.1016/j.ssaho.2025.101933>
- Karunanayake, K. S. (2022). Sources of second language speaking anxiety: An investigation into the lecturers' perspective. *International Journal of KIU*, 3(2), 90–97. <https://ij.kiu.ac.lk/article/read/47>



## Original Article

- Krashen, S. D., & Terrell, T. D. (1983). *The natural approach: Language acquisition in the classroom*. Pergamon Press.
- Long, M. H. (1983). Native speaker/non-native speaker conversation and the negotiation of comprehensible input. *Applied Linguistics*, 4(2), 126–141.
- Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. Ritchie & T. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413–468). Academic Press.
- Mackey, A., & Gass, S. M. (2014). *Second language research: Methodology and design* (2nd ed.). Routledge.
- Nuñez, A. A. C., Nuñez, M. S. C., Pachay, J. F. Z. P. Z., & Bosquez, A. M. C. B. C. (2025). Using ChatGPT voice to improve speaking skills in English language learners. *Ciencia Latina Revista Científica Multidisciplinar*, 9(1), 7143–7161. [https://doi.org/10.37811/cl\\_rcm.v9i1.16390](https://doi.org/10.37811/cl_rcm.v9i1.16390)
- Pratiwi, N., Efendy, A. G., Rini, H. C., & Ahmed, N. A. (2024). Speaking practice using ChatGPT's voice conversation: A review on potentials and concerns. *Journal of Language Intelligence and Culture*, 6(1), 59–72. <https://doi.org/10.35719/jlic.v6i1.149>
- Pushpakumara, P. B. S. L. (2024). Peer learning as a catalyst for engagement: A study on diploma in English students in a Sri Lankan university. *Journal of Language Studies*, 7(2), 135–208. <https://doi.org/10.4038/jls.v7i2.23>
- Salsabil, A. D., Rakhmawati, L. A., & Febtwenesty, N. (2025). From silent learners to confident speakers: The effect of AI voice chat with ChatGPT on EFL speaking skills. *Journal of English Education*, 3(1), 38–50. <https://doi.org/10.61994/jee.v3i1.1137>
- Santoso, K. H., Al Faruq, M. S., Kurniawati, K., Sabar, Hidayati, S. P., & Aung, Y. M. M. (2025). Leveraging ChatGPT's advanced voice mode for enhancing oral-aural skills in higher education language learning. *iMProvement Journal*, 12(2), 115–124. <https://journal.unj.ac.id/unj/index.php/improvement/article/view/62614/21930>
- Shalihah, A. (2025). Enhancing spoken English: Using ChatGPT and AI tools for personalized pronunciation and fluency training. In *Revolutionizing academic writing and language learning with AI: An educational perspective* (pp. 194–208). Eureka Media Aksara. [https://www.researchgate.net/publication/392030462\\_REVOLUTIONIZING\\_ACADEMIC\\_WRITING\\_AND\\_LANGUAGE\\_LEARNING\\_WITH\\_AI\\_an\\_Educational\\_Perspective](https://www.researchgate.net/publication/392030462_REVOLUTIONIZING_ACADEMIC_WRITING_AND_LANGUAGE_LEARNING_WITH_AI_an_Educational_Perspective)
- Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), *Principle and practice in applied linguistics* (pp. 125–144). Oxford University Press.
- Swain, M. (2000). The output hypothesis and beyond: Mediating acquisition through collaborative dialogue. In J. P. Lantolf (Ed.), *Sociocultural theory and second language learning* (pp. 97–114). Oxford University Press.
- Swain, M. (2005). The output hypothesis: Theory and research. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 471–484). Routledge.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.