FOSTER JELT
Journal of English Language Teaching

Copyright © The Author(s)
Vol.7, No. 3, July 2026
***e*-ISSN: 2723-4126**
***p*-ISSN: 2776-8880**

TECHNOLOGY ACCEPTANCE OF CHATGPT VOICE MODE IN ENGLISH SPEAKING PRACTICE: EVIDENCE FROM INTERNATIONAL STUDENTS IN INDONESIA

Mufid Lahiria Permata Hati[1*], Juita Triana[2], Zulia Chasanah[3]
[1,2,3] Universitas Muhammadiyah Purworejo, Purworejo, Indonesia

haimufidlhriaa@gmail.com

ABSTRACT

Non-native English-speaking students in international academic environments often face difficulties in developing speaking skills due to limited opportunities for practice, low confidence, and communication anxiety. Although artificial intelligence (AI)-based tools have been increasingly used to support language learning, limited research has examined international students' perceptions of ChatGPT voice mode in non-English-speaking educational contexts through the Technology Acceptance Model (TAM). This study explores students' perceptions of ChatGPT voice mode for English speaking practice, focusing on perceived usefulness, perceived ease of use, behavioral intention, and challenges encountered during its use. A quantitative survey was conducted with ten non-native English-speaking undergraduate students in Indonesia using a 20-item Likert-scale questionnaire, and the data were analyzed descriptively. The results indicate that students generally perceive the feature positively in terms of usefulness and ease of use (M = 3.0–3.4) and demonstrate a willingness to continue using it (M = 3.4). Reported challenges are relatively limited (M = 2.1–2.8), mainly related to speech recognition, accent differences, and limitations in replicating human interaction. Overall, ChatGPT voice mode functions as a supportive tool for independent speaking practice. The findings contribute to research on AI-assisted language learning and provide practical insights for integrating AI-supported speaking tools into language learning environments.

Keywords: AI-assisted Language Learning; ChatGPT Voice Mode; English-Speaking Practice; International Students; Technology Acceptance Model

INTRODUCTION

As a lingua franca, English plays an important role in the international context of both spoken and written communication among speakers with different native languages (Tan, 2024). Furthermore, English serves as a bridging language that facilitates communication across linguistic and cultural boundaries (Mikasa et al., 2024; Rao, 2019). Li et al. (2024) & Tenenti (2023) report that there are more than 1.5 billion English language learners worldwide, accounting for over 20% of the global population and approximately four times the number of native English speakers. Furthermore, Laputková (2025) state that 55 countries in the world use English as their official language, representing about 29% of the world's population.

FOSTER JELT
Journal of English Language Teaching

Copyright © The Author(s)
Vol.7, No. 3, July 2026
***e*-ISSN**: **2723-4126**
***p*-ISSN**: **2776-8880**

Aligned with this global trend, English is becoming increasingly important in the international academic environment, especially in inclusive universities in Indonesia. In this context, international students, especially those from non-English-speaking countries, must be able to communicate effectively in English in order to support both academic and interpersonal interactions (Puspita et al., 2025). Achieving a high level of English proficiency requires students to master the four main language competencies, which are listening, speaking, reading, and writing (Pratiwi et al., 2023). Among these skills, speaking is considered one of the most essential abilities because it enables learners to engage in direct interaction and express their ideas in real-time communication (Zahra & Ifadah, 2025). In the context of formal speaking, students are required to organize complex ideas and provide feedback.

Speaking can be defined as one of the most complex activities in the English language learning process, as it functions as a medium for expressing ideas, conveying emotions, and delivering information to others (Afifah & Devana, 2020; Nazri, 2021). Brown (2004), as cited in Razi et al. (2021) also states that speaking is an interactive process of constructing meaning in which individuals produce, receive, and process information orally. In this process, speakers need to organize their thoughts and convey them clearly so that the intended message can be properly understood by the listener. In speaking activities, several linguistic components work together during communication. These components include vocabulary, grammar, pronunciation, and fluency (Satriawan & Skolastika, 2023).

(Akhter, 2021) state that it is essential for students to develop effective speaking skills both in their academic setting and in the wider context of the global community. This is due to the globalization of social interaction, which has led to more frequent communication between countries and cultures. Lubis and Wahono (2024) state that mastery of speaking skills is of utmost importance because speaking enables speakers to support and participate in international communication in various aspects of life, such as education, career, and cross-cultural interaction. The more proficient speakers are, the more freedom speakers have to navigate and thrive in the challenges of globalization. Furthermore, in an academic context, students are expected not only to learn English as a supplement to their report card grades but also to apply it in their lives. This allows students to engage in academic discussions through speaking and increase their international career opportunities (El-Haj et al., 2025).

Despite the importance of speaking skills, students need to understand that this ability can only be achieved if they have a good command of the linguistic components of speaking, sufficient opportunities, and awareness of foreign languages. A learner's inner readiness to speak plays an important role in encouraging individuals to participate in English speaking activities (Darasawang & Reinders, 2021; Kamalia & Salim, 2025). However, not all students are able to speak English fluently or willingly according to its components, due to various factors. Students sometimes have difficulty expressing themselves due to limited vocabulary, low self-confidence, anxiety, and fear of making mistakes (Baharuddin, 2020). These conditions indicate that speaking ability is about more than just linguistic mastery; it also involves psychological factors and a supportive learning environment.

FOSTER JELT
Journal of English Language Teaching

Copyright © The Author(s)
Vol.7, No. 3, July 2026
***e*-ISSN: 2723-4126**
***p*-ISSN: 2776-8880**

The development of conventional learning methods, where learners use books or interactions with others to acquire and memorize vocabulary, then apply it to contextual situations, becomes less than effective if learners consider this to be the only way to learn to speak English. The reason for this is that it requires a lengthy process, unequal opportunities, differences in the speaking abilities of each individual, and access to learning resources is limited (Mustamir, 2024; Wahyuningsih & Afandi, 2020). Technological developments have created a dynamic and fast-paced environment. Therefore, the involvement of technology such as Artificial Intelligence as part of contemporary learning methods is more relevant to the needs of learners in the 21st century (Tusino et al., 2025).

In this context, technological advancements are widely recognized as a key factor in supporting English language learning, particularly in fostering interactive and flexible learning environments. The effectiveness of digital tools is closely linked to meaningful user engagement, thereby enabling learners to actively participate in the learning process (Johan et al., 2026; Anggraeni et al., 2021; Sukarni et al., 2025). More specifically, in the context of speaking, interactive and communicative tools are considered beneficial in fostering students' confidence and engagement, as well as improving their communication performance. This approach demonstrates that technology serves not only as a support for structured learning activities but also for more independent speaking practice. Consequently, technology enables students to use language in a more active and interactive manner (Wijayanti & Chasanah, 2021; Shafira et al., 2025).

Building on this development, AI can be understood as the development of systems that mimic human thinking abilities, focusing on areas of computer science such as seeking and acquiring new knowledge, learning, problem-solving, and making decisions related to those problems (L. Chen et al., 2020; Nur, 2023). As part of the technological revolution in computing and information processing, AI has brought about significant changes across various fields such as healthcare, the automotive industry, and education. In the field of education, there is great potential for leveraging the benefits of AI, given the increasing integration of learning activities with advancements in internet technology. Tusino et al. (2025) state that this can be seen in artificial intelligence-based learning systems that support more personalized learning without relying on the number of teachers, known as Intelligent Tutoring Systems (ITS).

Artificial intelligence provides a more efficient approach to learning to speak. This is evident in its natural language-based interaction features, which allow users to engage in two-way communication (Artha et al., 2026). From a behaviorist perspective, AI is also capable of supporting self-directed learning by providing immediate feedback (Rabuandika, 2025). According to Kufel et al. (2023) AI technology operates by leveraging large amounts of data, which is then processed through artificial neural network-based learning systems. Through this process, AI is able to recognize patterns, understand information, and generate responses more effectively and efficiently. AI-based speaking practice allows users to improve their speaking skills through a derivative of AI known as a chatbot.

A Chatbot is an artificial intelligence innovation in computer programming that can process natural language, whether through text or voice, in order to mimic human conversation (Caldarini et al., 2022). As opposed to conventional digital devices, which

FOSTER JELT
Journal of English Language Teaching

Copyright © The Author(s)
Vol.7, No. 3, July 2026
***e*-ISSN: 2723-4126**
***p*-ISSN: 2776-8880**

generally only execute limited commands, this technology enables the system to understand questions or commands from users, thereby generating relevant and interactive responses. This capability is supported by chatbot systems that can automatically respond to conversations (Shilpa & Menon, 2024; Singh, 2025). With the advancement of artificial intelligence technology, the use of chatbots has also expanded across various fields, including education. In the context of speaking practice, chatbots can be utilized as a supportive tool to help students access information, practice, and receive feedback in a more flexible manner (Guamán & Ramírez, 2025; Duc, 2025; Mishra et al., 2018; Wu et al., 2026). Currently, one of the most widely used chatbots is ChatGPT, which was developed using the Generative Pre-trained Transformer (GPT) model and is capable of generating responses that resemble human conversation.

Ever since its launch in November 2022, ChatGPT has quickly garnered global attention and amassed over 100 million users within two months (Halaweh, 2023). The surge in popularity of this technology has also spurred the development of subscription-based services that offer additional features such as faster responses and priority access. In the field of education, ChatGPT has the potential to enhance language learning, particularly in improving speaking skills. Through interactive conversations, students can practice expressing their opinions, expand their vocabulary, and reduce the fear of making grammatical mistakes when communicating in English (Nuñez et al., 2025; Salsabil et al., 2026). This involvement can help students build confidence and improve their speaking skills more effectively.

The voice mode in ChatGPT represents a gradual transition from voice-based interactions, which are still limited, toward more natural and responsive conversations (Saptiany et al., 2025). Initially, this feature only allowed users to interact via voice, with sound quality that was not yet natural. With subsequent updates, the system has become better equipped to support smoother real-time dialogue and demonstrates greater flexibility across various usage scenarios. These advancements make voice mode increasingly relevant for language learning, particularly in the context of developing speaking skills and multimodal communication (Nguyet & Nhan, 2024).

Several studies have investigated the integration of ChatGPT and other artificial intelligence (AI)-based technologies in supporting English speaking development. Existing research generally demonstrates that AI-assisted learning contributes positively to students' speaking performance, confidence, and communication skills. For instance, Noe et al. (2025) found that ChatGPT functioned effectively as a speaking partner that reduced communication anxiety, fear of negative evaluation, and test-related stress. Similarly, Muzamil and Allahbaksh (2025) reported that AI-supported speaking activities improved learners' vocabulary development, pronunciation accuracy, sentence coherence, and interactive communication after several weeks of implementation. Çolak (2024) also revealed that integrating ChatGPT into repetitive speaking tasks enhanced students' grammatical accuracy, task completion, motivation, and overall attitudes toward English language learning. Furthermore, Widyawati (2026) systematic review highlighted that teachers perceived ChatGPT as a valuable instructional tool because it provides immediate feedback and facilitates more interactive speaking practice.

FOSTER JELT
Journal of English Language Teaching

Copyright © The Author(s)
Vol.7, No. 3, July 2026
***e*-ISSN**: **2723-4126**
***p*-ISSN**: **2776-8880**

Collectively, these studies indicate that ChatGPT has significant potential to support English speaking development through interactive communication, personalized feedback, and flexible learning opportunities. The findings also suggest that AI-based learning environments may help learners overcome psychological barriers such as anxiety and low self-confidence, which are commonly experienced in speaking activities. In addition, recent developments in ChatGPT voice mode have created more authentic opportunities for real-time oral interaction, allowing learners to engage in more natural communication practice compared to conventional text-based interaction.

However, despite the growing body of literature on ChatGPT-assisted language learning, several important gaps remain insufficiently explored. First, most previous studies have focused primarily on general EFL learners or English-major students within controlled classroom settings. As a result, limited attention has been given to international students studying in non-English-speaking environments, particularly those who rely on English as a medium for both academic and social communication. This context is important because international students often face unique linguistic, psychological, and intercultural communication challenges that differ from those experienced by regular EFL learners.

Second, previous studies have predominantly emphasized text-based interaction with ChatGPT, while research examining the educational potential of ChatGPT voice mode remains relatively limited. Although voice mode enables more authentic spoken interaction and immediate conversational feedback, its role in supporting speaking practice among non-native international students has not been extensively investigated. This indicates a need for further exploration regarding how learners perceive and utilize voice-based AI interaction in real speaking contexts.

Third, prior studies have mainly examined the effectiveness of ChatGPT in improving speaking performance, motivation, or anxiety reduction, while limited research has investigated students' acceptance of the technology itself using theoretical frameworks such as the Technology Acceptance Model (TAM). Understanding learners' perceptions regarding perceived usefulness, perceived ease of use, and behavioral intention is essential in determining whether AI-assisted speaking tools can be sustainably integrated into language learning practices.

To address these gaps, this study investigates non-native international students' perceptions of using ChatGPT voice mode for English speaking practice in Indonesia. Specifically, this research explores students' perceived usefulness, perceived ease of use, behavioral intention, and the challenges they encounter when using ChatGPT voice mode within the framework of the Technology Acceptance Model (TAM). By focusing on international students in a non-English-speaking environment and emphasizing voice-based interaction, this study seeks to contribute to the growing discussion on AI-assisted language learning and provide insights into the practical use of conversational AI for authentic speaking development. This research was conducted to answer the following questions:

To what extent do international students perceive the usefulness, ease of use, and behavioral intention of using ChatGPT's voice mode for English speaking practice in Indonesia?

FOSTER JELT
Journal of English Language Teaching

Copyright © The Author(s)
Vol.7, No. 3, July 2026
***e*-ISSN**: **2723-4126**
***p*-ISSN**: **2776-8880**

What challenges do these students encounter when using ChatGPT's voice mode for English speaking practice?

METHODS

This study employed a quantitative research design using a survey method to investigate international students' perceptions of using ChatGPT voice mode for English speaking practice. The quantitative approach was considered appropriate because it enables the systematic measurement and analysis of the key constructs of the Technology Acceptance Model (TAM), namely perceived usefulness, perceived ease of use, and behavioral intention. In addition, the survey method allows researchers to collect structured data efficiently from a specific participant group and identify trends in students' perceptions through descriptive statistical analysis (Wasosa, 2025; Oranga & Matere, 2025).

The participants of this study consisted of ten international undergraduate students studying at a Muhammadiyah university in Purworejo, Central Java, Indonesia. The participants included five students from Mali and five students from Timor-Leste, all of whom came from countries where English is not used as an official language. They were enrolled in non-language-related programs, including Computer Science, Management, Law, and Animal Science, and relied on English for both academic and social communication.

Participants were selected using purposive sampling based on three criteria: (1) being international students from non-English-speaking backgrounds, (2) not majoring in English-related programs, and (3) having prior experience using ChatGPT voice mode for speaking practice. These criteria were intentionally established to ensure that participants were relevant to the objectives of the study and capable of providing meaningful insights into the practical use of ChatGPT voice mode in authentic communication contexts.

All participants were between 19 and 21 years old and demonstrated varying levels of engagement with ChatGPT voice mode, ranging from occasional to frequent use. Although the sample size was relatively small, it was considered appropriate for the exploratory nature of the study, which focused on obtaining initial insights from a specific participant group. In addition, the limited number of international students who met the study criteria influenced the sample size. Nevertheless, this limitation should be acknowledged, as the findings may not fully represent the experiences of international students in different educational contexts.

The primary instrument used in this study was a closed-ended questionnaire consisting of twenty items developed based on the constructs of the Technology Acceptance Model (TAM) proposed by Davis (1989)(Davis & Granic, 2024), particularly perceived usefulness, perceived ease of use, and behavioral intention. The questionnaire items were adapted from previous TAM-related studies on educational technology acceptance to ensure conceptual relevance and consistency. In addition, items related to challenges in using ChatGPT voice mode for speaking practice were developed based on previous studies discussing AI-assisted language learning and chatbot interaction. Prior to data collection, the questionnaire items were reviewed to ensure clarity, appropriateness, and alignment with the objectives of the study.

FOSTER JELT
Journal of English Language Teaching

Copyright © The Author(s)
Vol.7, No. 3, July 2026
***e*-ISSN: 2723-4126**
***p*-ISSN: 2776-8880**

The questionnaire was administered through Google Forms and distributed via online communication platforms to facilitate accessibility and flexibility for participants. Participants were asked to indicate their level of agreement with each statement based on their experiences using ChatGPT voice mode. After all responses were collected, the data were systematically organized, coded, and analyzed using descriptive statistical analysis to identify patterns in students' perceptions regarding the usefulness, ease of use, behavioral intention, and challenges associated with ChatGPT voice mode.

Through this analysis, the study provides an overview of how international students perceive the effectiveness of ChatGPT voice mode in supporting speaking fluency, confidence, and interactive communication practice. In addition, the analysis identifies several challenges experienced by participants, including limitations in speech recognition accuracy, sensitivity to diverse accents, and limited interaction depth compared to human communication.

RESULTS

After analyzing 10 international students for whom English is not their native language, their perceptions regarding how to use ChatGPT's voice mode for speaking practice were identified. This study used Google Forms as a data collection tool via a questionnaire. From the 10 responses received, the researcher collected and analyzed the students' answers, which are presented in the following tables

Table 1. Questionnaire results

No	Statement	Strongly Agree	Agree	Disagree	Strongly Disagree
1.	ChatGPT voice mode enhances students' English communication	40 %	40%	20 %	0%
2.	ChatGPT voice mode improves students' speaking skills.	30 %	70%	0%	0%
3.	ChatGPT voice mode supports real-life communication practice	40 %	50 %	10 %	0%
4.	ChatGPT voice mode increases speaking confidence.	30%	60%	10%	0%
5.	ChatGPT voice mode provides useful speaking feedback	30%	60%	10%	0%
6.	ChatGPT voice mode is easy to use for practicing English communication.	40%	50%	10%	0%
7.	ChatGPT voice mode is easy to operate.	40%	50%	10%	0%
8.	Students interact with ChatGPT voice mode easily.	20%	60%	20%	0%
9.	ChatGPT voice mode fits daily learning routines.	20%	70%	20%	0%
10.	ChatGPT voice mode is easy to learn.	10%	70%	0%	10%

FOSTER JELT
Journal of English Language Teaching

Copyright © The Author(s)
Vol.7, No. 3, July 2026
***e*-ISSN: 2723-4126**
***p*-ISSN: 2776-8880**

11.	Students intend to continue using ChatGPT voice mode.	20%	70%	10%	0%
12.	Students plan to use ChatGPT voice mode regularly.	10%	50%	40%	0%
13.	Students recommend ChatGPT voice mode to others.	20%	80%	0%	0%
14.	Students prefer ChatGPT voice mode over other tools	20%	30%	40%	10%
15.	Students face difficulties using ChatGPT voice mode.	0%	20%	70%	10%
16.	ChatGPT voice mode misrecognizes speech.	10%	60%	20%	10%
17.	Challenges Students' accents affect recognition by ChatGPT voice mode.	0%	80%	20%	0%
18.	Technical issues affect ChatGPT voice mode use.	0 %	70%	30%	0%
19.	Student feel uncomfortable to talk with ChatGPT voice mode.	10%	10%	80%	0 %
20.	The ChatGPT voice mode cannot replace human interaction.	30%	40%	20%	0%

Table 1 shows that the majority of responses fall into the "agree" and "strongly agree" categories, particularly for items related to Perceived Usefulness (PU) and Perceived Ease of Use (PEOU). This indicates that students generally consider the feature to be useful and relatively easy to use. Meanwhile, in the Behavioral Intention (BI) category, students showed a tendency to continue using and recommending the feature. However, lower levels of agreement were observed on items related to routine use and preference for other tools, indicating variation in long-term usage intentions.

Furthermore, in the Challenges dimension, the majority of students reported that they did not experience significant difficulties or discomfort when using the feature. At the same time, a higher proportion of students acknowledged issues related to voice recognition, the influence of accents, and the feature's limitations in replacing human interaction. Overall, the results indicate a generally positive perception of ChatGPT's voice mode feature, with some variation in long-term usage intentions and technical challenges.

Furthermore, to provide a clearer numerical summary, the average score for each item was calculated and presented in Table 2. These values indicate the general trend in students' perceptions across each dimension and facilitate comparisons between categories.

Table 2. Perceived usefulness (PU) of ChatGPT voice mode

Perception	Statement Items	Mean
Perceived Usefulness	ChatGPT voice mode enhances students' English communication	3
	ChatGPT voice mode improves students' speaking skills.	3.3
	ChatGPT voice mode supports real-life communication practice.	3.3

FOSTER JELT
Journal of English Language Teaching

Copyright © The Author(s)
Vol.7, No. 3, July 2026
***e*-ISSN: 2723-4126**
***p*-ISSN: 2776-8880**

	ChatGPT voice mode increases speaking confidence.	3
	ChatGPT voice mode provides useful speaking feedback.	3.2

Table presents the results for the Perceived Usefulness (PU) variable. The average scores for all questions fell within the positive range of 3 to 3.3. Among these scores, the highest average (M = 3.3) was found in items related to improved speaking skills and support for communication in daily life. Meanwhile, an average score of 3 was shown by items measuring English communication skills and confidence in speaking. Furthermore, the feedback variable received an average score of 3.2. Thus, all PU variables fall within the positive category.

Table 3. Perceived Ease of Use (PEOU) of ChatGPT Voice Mode

Perception	Statement Items	Mean
Perceived Ease of Use (PEOU)	ChatGPT voice mode is easy to use for practicing English communication.	3.3
	ChatGPT voice mode is easy to operate.	3.4
	Students interact with ChatGPT voice mode easily.	3
	ChatGPT voice mode fits daily learning routines.	3.4
	ChatGPT voice mode feature is easy to learn.	3

Table 3 shows the Perceived Ease of Use (PEOU) as reported by the respondents. In general, the students found ChatGPT's voice mode to be fairly easy to use, with the positive category scoring an average between 3.0 and 3.4. The highest average score was obtained for ease of operation (M = 3.4), indicating that students were able to interact with this feature without encountering significant difficulties. In other aspects, such as ease of use and ease of learning, the scores were relatively similar, ranging from (M = 3.0–3.3), reflecting a consistent level of usability across various indicators. Additionally, a slightly higher perception was shown regarding the alignment of these features with students' daily learning routines (M = 3.2), illustrating their relevance to routine learning activities.

Table 4. Behavioral intention (BI) toward ChatGPT voice mode

Perception	Statement Items	Mean
Behavioral Intention (BI)	Students intend to continue using ChatGPT voice mode.	3.4
	Students plan to use ChatGPT voice mode regularly.	2.2
	Students recommend ChatGPT voice mode to others.	3.2
	Students prefer ChatGPT voice modeover other tools.	2.4

Next, Table 4 presents data on students' Behavioral Intention (BI) regarding the use of ChatGPT's voice mode. The data show that the average scores range from 2.4 to 3.4. The highest score was found in students' intention to continue using the feature (M = 3.4), which falls into the positive category. This is followed by students' intention to recommend the feature to others (M = 3.2), which also falls into the positive category. Meanwhile, a lower score was found in students' intention to use the feature regularly (M = 2.4), which falls into the negative category. Furthermore,

FOSTER JELT
Journal of English Language Teaching

Copyright © The Author(s)
Vol.7, No. 3, July 2026
***e*-ISSN: 2723-4126**
***p*-ISSN: 2776-8880**

students' preference for this feature compared to other tools also falls into the negative category (M = 2.4).

Table 5. Perceived Challenges in Using ChatGPT Voice Mode

Perception	Statement Items	Mean
Challenges	Students face difficulties using ChatGPT voice mode	2.1
	ChatGPT voice mode misrecognizes speech.	2.7
	Students' accents affect recognition by ChatGPT voice mode	2.8
	Technical issues affect ChatGPT Voice mode use	2.7
	Student feel uncomfortable to talk with ChatGPT voice mode.	2.3
	The ChatGPT voice mode cannot replace human interaction.	2.8

Table 5 presents the results for the Challenges variable, with mean scores ranging from 2.1 to 2.8, indicating a generally low level of perceived difficulty. The lowest score (M = 2.1) appears in students' difficulty in using the ChatGPT voice mode, suggesting that most participants did not find it particularly hard to use. Scores related to speech recognition errors and technical issues are slightly higher, both at 2.7, showing that these aspects are experienced to a certain extent. The highest scores (M = 2.8) are found in the influence of students' accents on recognition and the view that the feature cannot replace human interaction. Meanwhile, students' discomfort when using the feature remains relatively low (M = 2.3). Overall, while some challenges are reported, they do not seem to significantly hinder students in using the feature.

DISCUSSION

This study examines how non-English-speaking international students' perceptions of the use of ChatGPT's voice mode influence their speaking ability. The study found that the use of this feature has a significant impact on students' speaking ability, both positively and negatively. The researcher categorized the students' perceptions into two categories. The first category consists of Perceived Usefulness, Perceived Ease of Use, and Behavioral Intention to Use the ChatGPT Voice Mode. The second category focuses on the challenges faced when using the ChatGPT voice mode.

Perceived Usefulness, Perceived Ease of Use, and Behavioral Intention to Use ChatGPT Voice Mode

The findings indicate that international students generally perceived ChatGPT voice mode as beneficial for supporting English speaking practice, particularly in improving speaking skills, communication practice, speaking confidence, and speaking feedback. The high scores in these indicators suggest that voice-based AI interaction may provide more flexible opportunities for authentic oral communication, especially for students studying in non-English-speaking environments where access to English-speaking interaction is

FOSTER JELT
Journal of English Language Teaching

Copyright © The Author(s)
Vol.7, No. 3, July 2026
***e*-ISSN**: **2723-4126**
***p*-ISSN**: **2776-8880**

relatively limited. In this context, ChatGPT voice mode may function not only as a technological tool but also as a supplementary communication partner that enables students to practice speaking more independently and consistently.

This finding supports the Technology Acceptance Model (TAM), which emphasizes that users are more likely to adopt a technology when they perceive it as useful for improving their performance and achieving their learning goals (Davis, 1989). The positive perception toward speaking improvement and communication practice also indicates that students recognize the practical value of AI-assisted speaking tools in facilitating real-time interaction and immediate feedback. This result is consistent with previous studies showing that AI-based speaking tools can enhance students' communication confidence, speaking fluency, and learning motivation (Nugroho et al., 2024). Furthermore, the findings may indicate that international students tend to value learning tools that provide flexible and low-pressure speaking environments. Unlike classroom interaction, ChatGPT voice mode allows students to practice speaking without fear of direct judgment from teachers or peers. As a result, the feature may help reduce communication anxiety and encourage students to participate more actively in speaking activities.

The findings on perceived ease of use demonstrate that students generally considered ChatGPT voice mode easy to operate and integrate into their daily learning activities. This suggests that the accessibility and simplicity of the feature may reduce technological barriers that often discourage students from using digital learning tools independently. For international students who are already adapting to new academic and social environments, ease of use becomes particularly important because complicated systems may increase cognitive burden and reduce learning motivation. From the perspective of the Technology Acceptance Model (TAM), this finding reinforces the idea that perceived ease of use plays an important role in shaping students' willingness to adopt educational technology. When students perceive a technology as simple and manageable, they are more likely to use it continuously and integrate it into their learning routines. This finding is also supported by Cognitive Load Theory, which suggests that simpler systems allow learners to focus more on meaningful learning tasks rather than technical difficulties (Sweller, 1988). Although students generally perceived the feature positively, slight variations in scores suggest that some participants still required adjustment in becoming familiar with voice-based AI interaction. This indicates that while ChatGPT voice mode is relatively accessible, students' technological familiarity and speaking habits may still influence their learning experience.

The findings on behavioral intention reveal that students generally expressed willingness to continue using and recommending ChatGPT voice mode for speaking practice. This indicates that students recognize the practical benefits of the feature and consider it supportive for independent language learning. In the context of TAM, positive behavioral intention reflects students' acceptance of the technology as a potentially valuable learning tool. However, lower scores related to regular use and preference over other learning tools suggest that students do not perceive ChatGPT voice mode as a complete replacement for

FOSTER JELT
Journal of English Language Teaching

Copyright © The Author(s)
Vol.7, No. 3, July 2026
***e*-ISSN: 2723-4126**
***p*-ISSN: 2776-8880**

conventional learning methods. Instead, students appear to position the feature as a complementary learning tool that supports, rather than substitutes, human interaction and formal instruction. This finding is important because it highlights that technology acceptance does not necessarily indicate full dependency on AI-assisted learning systems.

The moderate intention to use the feature regularly may also reflect students' awareness of the limitations of AI-based interaction, particularly in terms of emotional communication, contextual understanding, and interpersonal feedback. Therefore, while ChatGPT voice mode is positively accepted, students still value direct interaction with teachers and peers as an essential aspect of language learning. This finding is consistent with previous studies indicating that learners continue to prefer human feedback and social interaction in language learning despite the growing use of AI-assisted technologies (Chen, 2025; Fang & Han, 2025).

Challenges in Using ChatGPT Voice Mode for English Speaking Practice

Regarding the challenges encountered, the findings show that students generally did not experience major difficulties or discomfort when using ChatGPT voice mode. This indicates that the feature is relatively accessible and manageable for international students despite their non-native English backgrounds. Nevertheless, several technical and interactional limitations were still identified, particularly related to speech recognition accuracy, accent sensitivity, and occasional technical issues. These findings suggest that although AI-based speaking tools offer flexibility and accessibility, their effectiveness may still be influenced by linguistic diversity and technological limitations in real communication contexts.

A more prominent challenge identified in this study relates to the influence of students' accents on speech recognition performance. Variations in pronunciation may affect the system's ability to accurately interpret spoken input, especially for users whose accents differ from the dominant recognition patterns used in AI training systems. This finding aligns with previous studies indicating that non-native speakers often experience greater challenges when interacting with automatic speech recognition systems due to accent variation (Banerjee & Ramasubramanian, 2025; Radzikowski, 2021; Tong et al., 2022). Although these limitations do not significantly reduce the overall usefulness of the feature, they demonstrate that AI-assisted speaking tools may not yet provide equally accurate interaction experiences for all learners.

In addition, students tended to perceive that ChatGPT voice mode could not fully replace human interaction in language learning. This finding reflects the social and interpersonal dimensions of speaking practice, where effective communication involves not only linguistic accuracy but also emotional, cultural, and contextual understanding. While ChatGPT voice mode can facilitate interactive speaking practice, it still lacks the nuanced feedback and interpersonal responsiveness typically found in human communication. Therefore, the findings suggest that ChatGPT voice mode is better understood as a

FOSTER JELT
Journal of English Language Teaching

Copyright © The Author(s)
Vol.7, No. 3, July 2026
***e*-ISSN: 2723-4126**
***p*-ISSN: 2776-8880**

complementary tool rather than a substitute for teachers, peers, or authentic social interaction in language learning contexts.

CONCLUSION

This study investigated non-English-speaking international students' perceptions toward the use of ChatGPT voice mode for English speaking practice. The findings indicate that participants generally perceived the feature as useful and easy to use, particularly in supporting speaking confidence, communication practice, and independent learning. These findings support the Technology Acceptance Model (TAM), which emphasizes that perceived usefulness and perceived ease of use influence users' acceptance of technology. The results further suggest that when a learning tool is accessible, flexible, and beneficial, learners are more likely to engage with it as part of their learning process. In addition, ChatGPT voice mode appears to provide a supportive environment for speaking practice, especially for international students who may have limited opportunities for authentic English communication in non-English-speaking settings.

However, although participants showed positive perceptions toward ChatGPT voice mode, they did not demonstrate a strong intention to use it regularly or rely on it as their primary learning tool. This suggests that AI-based speaking tools are still perceived as complementary learning support rather than substitutes for direct interaction with teachers or peers. Several challenges were also identified, including speech recognition inaccuracies, accent sensitivity, and technical limitations during voice interaction, although these issues did not significantly reduce participants' overall positive perceptions of the feature. This study contributes to the growing discussion of AI-assisted language learning, particularly regarding the use of voice-based AI interaction in non-English-speaking educational contexts. Nevertheless, the study was limited by the relatively small number of participants and its focus on a single educational setting, which may affect the generalizability of the findings. In addition, the study focused only on learners' perceptions and did not directly measure actual speaking performance. Therefore, future research is recommended to involve larger and more diverse participant groups, apply mixed-method or experimental approaches, and further examine the long-term impact of ChatGPT voice mode on speaking development and learner engagement.

ACKNOWLEDGMENT

Acknowledgments are extended to Juita Triana, M.Pd. and Zulia Chasanah, S.S., M.Pd., whose supervision and constructive guidance have been essential throughout the completion of this study. Appreciation is also directed to Andrian Nuriza Johan, S.S., M.Pd., Head of the English Education Study Program, for his academic direction and consistent availability for consultation during the research process, while gratitude is further extended to the international students of Universitas Muhammadiyah Purworejo, particularly those from Mali and Timor Leste, whose participation as respondents provided valuable input for this study. Their contributions are duly acknowledged.

FOSTER JELT
Journal of English Language Teaching

Copyright © The Author(s)
Vol.7, No. 3, July 2026
***e*-ISSN**: **2723-4126**
***p*-ISSN**: **2776-8880**

REFERENCES

Abu El-Haj, S. A., Abu El-Haj, H., & Al-Assaf, N. M. (2025). English learning perceptions and career implications: insights from tertiary-level students. *International Journal of Evaluation and Research in Education*, *14*(4), 3325–3335. https://doi.org/10.11591/ijere.v14i4.32836

Afifah, N., & Devana, T. (2020). Speaking Skill through Task Based Learning in English Foreign Language Classroom. *Jo-ELT (Journal of English Language Teaching) Fakultas Pendidikan Bahasa & Seni Prodi Pendidikan Bahasa Inggris IKIP*, *7*(2), 135. https://doi.org/10.33394/jo-elt.v7i2.3109

Akhter, S. (2021). Exploring The Significance Of Speaking Skill For Efl Learners. *Journal of Archaeology of Egypt / Egyptology, 17(9), 6019–6030.*, *4*(3), 1–9. https://archives.palarch.nl/index.php/jae/article/view/5149

Albl-Mikasa, M., Ehrensberger-Dow, M., Gieshoff, A. C., & Hunziker Heeb, A. (2024). English as a lingua franca in interpreting and translation: a survey of practitioners. *Frontiers in Communication*, *9*(July), 1–13. https://doi.org/10.3389/fcomm.2024.1413110

Artha, B., Santoso, F., Sari, U. T., Bahri, B., Sari, N. P., Hadi, A. S., Asri, C. P., & Aditya, A. (2026). *A Narrative Review of English Language Learning and Artificial Intelligence*. *06*(1), 47–56.

Baharuddin, N. N. K. (2020). Efl Student Problems in Speaking : Case on Students of. *Jurnal Akademik FKIP UNDAYAN*, (124), 102–113.

Banerjee, T., & Ramasubramanian, V. (2025). *Accent - robust speech recognition for English in low - resource settings using Manifold Mixup*. *0*.

Caldarini, G., Jaf, S., & McGarry, K. (2022). A Literature Survey of Recent Advances in Chatbots. *Information (Switzerland)*, *13*(1). https://doi.org/10.3390/info13010041

Carrera Nuñez, A. A., Carrera Nuñez, M. S., Zambrano Pachay, J. F. Z. P., & Carrera Bosquez, A. M. C. B. (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

Chen, L., Chen, P., & Lin, Z. (2020). Artificial Intelligence in Education: A Review. *IEEE Access*, *8*, 75264–75278. https://doi.org/10.1109/ACCESS.2020.2988510

Chen, M. (2025). *The human touch in AI : optimizing language learning through self-determination theory and teacher scaffolding*. (July), 1–18. https://doi.org/10.3389/fpsyg.2025.1568239

Çolak, M. K. (2024). Focus on ELT journal enhancing speaking skills through task repetition and ChatGPT integration in remedial EFL lessons : An action research approach. *Focus on ELT Journal*, *6*(4), 1–6.

Darasawang, P., & Reinders, H. (2021). Willingness to communicate and second language

FOSTER JELT
Journal of English Language Teaching

Copyright © The Author(s)
Vol.7, No. 3, July 2026
***e*-ISSN**: **2723-4126**
***p*-ISSN**: **2776-8880**

proficiency: A correlational study. *Education Sciences*, *11*(9). https://doi.org/10.3390/educsci11090517

Davis, F. D., & Granic, A. (2024). Book Review : The Technology Acceptance Model - 30 Years of TAM The background to TAM The Evolution , Revolution and Future of TAM. *Journal of University Teaching & Learning Practice*, *21*(2), 1–6.

Duc, L. T. K. D. (2025). *Exploring Students' Perceptions Of The Effectiveness Of Using Artificial Intelligence Chatbots For English Speaking Practice*. *231*(04), 29–36. https://doi.org/https://doi.org/10.34238/tnu-jst.13224

Fang, S., & Han, Z. (2025). *On the nascency of ChatGPT in foreign language teaching and learning*. 148–178. https://doi.org/10.1017/S026719052510010X

Galuh Kirana Wijayanti, & Chasanah, Z. (2021). Speech Functions by the Main Character in Wonder Movie in Teaching Speaking. *Scripta : English Department Journal*, *8*(2), 28–34. https://doi.org/10.37729/scripta.v8i2.778

Guartán Guamán, S. P., & Valdiviezo Ramírez, E. A. (2025). Using chatbots as a tool to personalize english as a foreign language learning: A literature review. *Runas. Journal of Education and Culture*, *6*(12), e250283. https://doi.org/10.46652/runas.v6i12.283

Halaweh, M. (2023). ChatGPT in education: Strategies for responsible implementation. *Contemporary Educational Technology*, *15*(2). https://doi.org/10.30935/cedtech/13036

Henry Wasosa. (2025). Influence of Psychological Well-Being and School Factors on Delinquency , During the Covid-19 Period Among Secondary School Students in Selected Schools in Nakuru County : Kenya. *International Journal of Research and Innovation in Social Science (IJRISS)*, *VII*(2454), 1175–1189. https://doi.org/10.47772/IJRISS

Johan, A. N., Chasanah, Z., & Purworejo, U. M. (2026). *Enhancing EFL Students' Listening Comprehension through Preparation for the TOEFL Application*. *7*(1), 129–145. https://doi.org/https://doi.org/10.24256/foster-jelt.v7i1.329

Kamalia, A., & Salim, M. A. (2025). Implementation of Duolingo Application to Improve English Speaking Skills of Tribakti Islamic University Students : Classroom Action Study. *English Education, Linguistics, and Literature Journal*, *4*(2024), 144–152.

Kufel, J., Bargieł-Łączek, K., Kocot, S., Koźlik, M., Bartnikowska, W., Janik, M., Czogalik, Ł., Dudek, P., Magiera, M., Lis, A., Paszkiewicz, I., Nawrat, Z., Cebula, M., & Gruszczyńska, K. (2023). What Is Machine Learning, Artificial Neural Networks and Deep Learning?—Examples of Practical Applications in Medicine. *Diagnostics*, *13*(15). https://doi.org/10.3390/diagnostics13152582

Language, C. (2026). *Enhancing EFL Speaking Fluency and Confidence through ChatGPT Voice in Contextual Learning at Madrasah Aliyah*. *6*(1), 121–134.

Laputková, B. (2025). Speaking Skills Improvement at University Level English Lesson.

Copyright © The Author(s)
Vol.7, No. 3, July 2026
***e*-ISSN**: **2723-4126**
***p*-ISSN**: **2776-8880**

Innovare Journal of Education, *13*(4), 39–46. https://doi.org/10.22159/ijoe.2025v13i4.55436

Li, H., Meng, L., Mu, K., & Wang, S. (2024). English language requirement and educational inequality: Evidence from 16 million college applicants in China. *Journal of Development Economics*, *168*. https://doi.org/10.1016/j.jdeveco.2024.103271

Lubis, D. N., & Wahono, S. S. (2024). Enhancing Efl Speaking Skills Through Spotify: the Power of Interactive Comments. *English Review: Journal of English Education*, *12*(2), 685–694. https://doi.org/10.25134/erjee.v12i2.9834

Mishra, S., Jain, S., & Kumar, H. (2018). Exploring the impact of AI chatbots on student engagement and motivation: A systematic review. *Computers in Human Behavior*, *88*(1), 224–233.

Mustamir, M. (2024). Indonesian Efl Learners and Speaking Anxiety: Insights From a Meta-Synthetic Analysis. *English Review: Journal of English Education*, *12*(2), 509–518. https://doi.org/10.25134/erjee.v12i2.9950

Muzamil, M., & Allahbaksh, S. (2025). *Exploring the Effectiveness of Chatgpt in Improving the*. *8*(4), 1440–1445.

Nazri, M. A. (2021). Communication Barriers in Speaking: Challenges Faced by Indonesian EFL Learners. *Print) Journal of English Language and Education*, *10*(5), 2025.

Nguyet, N. T. M., & Nhan, P. T. T. (2024). Voice Mode of ChatGPT: A New Support Tool for Teaching English in University Programs. *EPRA International Journal of Research and Development (IJRD)*, *9*(7), 140–143. https://doi.org/10.36713/epra2016

Noe, S., Yusuf, F. N., & Wirza, Y. (2025). *ChatGPT as a Speaking Partner : Investigating Its Impact on Students ' Speaking Anxiety*. *XI*(2), 190–199.

Nugroho, T. B., Riana, N. A., & Syahid, A. (2024). Perceptions of using ChatGPT's real-time feedback to enhance EFL speaking skills. *PESHUM: Jurnal Pendidikan, Sosial Dan Humaniora*, *4*(1), 842–850.

Nur, F. T. (2023). ELT FORUM 12(1) (2023) Journal of English Language Teaching Artificial intelligence (AI) technology in OpenAI ChatGPT application: A review of ChatGPT in writing English essay. *Journal of English Language Teaching*, *12*(1), 44–58. http://journal.unnes.ac.id/sju/index.php/elt

Oranga, J., & Matere, A. (2025). Quantitative Research: Types, Advantages, Generalizability&Limitations. *OALib*, *12*(09), 1–9. https://doi.org/10.4236/oalib.1113694

Pratiwi, A., Jabu, B., & Talib, A. (2023). Using Story Telling to Improve the Speaking Ability of the Third Grade Students of SMP Negeri 2 Pangkajene. *Performance: Journal of English Education and Literature*, *2*(4), 542. https://doi.org/10.26858/performance.v2i4.54289

FOSTER JELT
Journal of English Language Teaching

Copyright © The Author(s)
Vol.7, No. 3, July 2026
***e*-ISSN**: **2723-4126**
***p*-ISSN**: **2776-8880**

Puspita, I. T., Silvhiany, S., & Ismail, P. (2025). Higher Education Internationalization and English Language Practices: Academics Perspectives. *Edukasi: Jurnal Pendidikan Dan Pengajaranedukasi: Jurnal Pendidikan Dan Pengajaran*, *12*(1), 539–553. https://doi.org/10.19109/npj84575

Rabuandika, A. (2025). Leveraging AI in Self-Directed Learning: A Phenomenological Study of Master's Students' Experiences. *JP (Jurnal Pendidikan): Teori Dan Praktik*, *10*(1), 1–13. https://doi.org/10.26740/jp.v10n1.p1-13

Radzikowski, K. (2021). *Accent modification for speech recognition of non-native speakers using neural style transfer*.

Rahma Novia Anggraeni, Ngafif, A., & Chasanah, Z. (2021). The Effectiveness of Using Visual Dictionary to Teach Vocabulary at the Grade Students. *Scripta : English Department Journal*, *8*(2), 11–19. https://doi.org/10.37729/scripta.v8i2.1196

Rao, P. S. (2019). The role of English as a global language. *Reseasch Journal of English (RJOE)*, *4*(1), 65–79. www.rjoe.org.in

Razi, F., Muslem, A., & Fitrisia, D. (2021). Universitas Syiah Kuala, Banda Aceh, Indonesia. *English Education Journal (EEJ)*, *12(4)*(October 2021), 540–557. https://doi.org/10.24815/eej.v12i3.19136

Saptiany, S. G., Hartono, R., Mujiyanto, J., & Rustipa, K. (2025). Exploring the Role of ChatGPT Voice Mode in Reducing Foreign Language Speaking Anxiety Among Indonesian ESP Students. *Educational Process International Journal*, *19*(1). https://doi.org/10.22521/edupij.2025.19.554

Satriawan, K., & Skolastika, I. M. (2023). The analysis of students' speaking ability in speech at English literature department of mahasaraswati denpasar university. *BULLET : Jurnal Multidisiplin Ilmu*, *2*(3), 739–746. https://journal.mediapublikasi.id/index.php/bullet

Shafira, S. E., Rokhayati, T., & Chasanah, Z. (2025). Exploring Students' Perceptions of Narrative Role-Play Videos in Speaking Assessment. *Journal of English Language and Education*, *10*(3), 263–274. https://doi.org/10.31004/jele.v10i3.864

Shilpa, K., & Menon, D. (2024). Consumer's interaction with ChatGPT- a UTAUT perspective. *Science Talks*, *9*(October 2023), 100281. https://doi.org/10.1016/j.sctalk.2023.100281

Singh, H. (2025). AI-Powered Chatbots Transforming Customer Support through Personalized and Automated Interactions. *SSRN Electronic Journal*, *X*(X), 450–468. https://doi.org/10.2139/ssrn.5267858

Sukarni, S., Chasanah, Z., Maskuri, E. S., Laksono, P., & Anua Jah, N. J. (2025). TPACK Competence and Technology Use Among in-Service EFL Teachers. *Journal of Languages and Language Teaching*, *13*(2), 1027. https://doi.org/10.33394/jollt.v13i2.13833

TENENTI, A. (2023). The English. *Piracy and the Decline of Venice 1580 - 1615*, 56–86.

FOSTER JELT
Journal of English Language Teaching

Copyright © The Author(s)
Vol.7, No. 3, July 2026
***e*-ISSN**: **2723-4126**
***p*-ISSN**: **2776-8880**

https://doi.org/10.2307/jj.8085283.10

Tong, F., Li, T., Liao, D., Xia, S., Li, S., & Hong, Q. (2022). *applied sciences The XMUSPEECH System for Accented English Automatic Speech Recognition*. 0–7.

Tusino, T., Dewi, P., Hati, M. L. P., & Iglesia, M. J. C. (2025). ChatGPT in Academic Writing Classes: Exploring EFL Students' Perspectives. *Journal of Languages and Language Teaching*, *13*(4), 2162–2172. https://doi.org/10.33394/jollt.v13i4.15293

Wahyuningsih, S., & Afandi, M. (2020). Investigating English speaking problems: Implications for speaking curriculum development in Indonesia. *European Journal of Educational Research*, *9*(3), 967–977. https://doi.org/10.12973/EU-JER.9.3.967

Widyawati, A. (2026). *A Systematic Review o f Teachers ' Perspectives on Using Chat GPT to Teach Speaking : Benefits and Challenges*. *9*(1), 261–272.

Wu, T. T., Hapsari, I. P., & Huang, Y. M. (2026). Intelligent chatbot–supported collaborative learning: Impact on student engagement and English speaking skills. *ReCALL*, 1–19. https://doi.org/10.1017/S0958344025100451

Zahra, O. A., & Ifadah, M. (2025). Analysing EFL Learners Speaking Skills through Authentic Materials and Role Plays Activities Analysing EFL Learners Speaking Skills through Authentic Materials and Role Play Activities. *English Language and Literature International Conference (ELLiC) Proceedings*, *8*(1), 624–633. https://jurnal.unimus.ac.id/index.php/ELLIC/index