



Using ChatGPT as a virtual speaking tutor to boost EFL learners' speaking self-efficacy

BUNYAMIN CELIK  

YUNUS YILDIZ  

SABAN KARA  

English Language Teaching Department, Tishk International University, Erbil-KRI, Iraq

Abstract

Expressing thoughts and feelings efficiently is fundamental in daily, academic, and professional life. Accordingly, self-efficacy beliefs play pivotal roles in shaping learners' speaking performance through various dimensions. Higher education institutions assume responsibility for boosting students' speaking self-efficacy, thereby contributing to their self-confidence in different venues. Meanwhile, AI-integrated English classes are on the rise thanks to their overt advantages. In this regard, this study investigated the role of ChatGPT in improving speaking self-efficacy of EFL learners who were selected randomly among sophomore students in Advanced English classes at the tertiary level in Erbil, Iraq. Forty-four students whose ages ranged from 18 to 27 joined the study to represent the control or experimental group voluntarily within 8 weeks in the 2024–2025 academic year. Control group students took part in speaking sessions in classes, whereas experimental group students took advantage of ChatGPT as a speaking tutor to support their learning in classes with given topics in line with the coursebook. Data collection was implemented through speaking exams, questionnaires, and interviews, the analyses of which were conducted using SPSS 29 and MAXQDA. Findings unearthed that ChatGPT is a promising tool to increase students' speaking self-efficacy scores because it offers a welcoming atmosphere to receive constructive feedback, play recordings several times, and balanced instruction according to learners' levels. Thus, they could relieve speaking anxiety and speak more self-assuredly. The current study's findings may have varied reflections for interested parties engaged in integrating AI tools into speaking classes harmoniously.

Keywords: AI-integrated English; ChatGPT; reduced speaking anxiety; speaking self-efficacy

Introduction

English has emerged as a preeminent worldwide language, functioning as a crucial communication medium across various fields like diplomacy, tourism, economy, and education. The significance of language is due to its capacity to adapt, evolve, and change, therefore achieving worldwide recognition (Issa, 2024). Accordingly, English occupies a vital position in educational institutions to tailor to the specific needs of students (Yıldız, 2024). In this regard, numerous institutions provide English-oriented education to teach English professionally and extend its impact to other subjects including sociology, history, geography, computer, and biology. In addition to encompassing courses in English for various fields, English occupies significant time in curricula to enhance learners' linguistic competence. Proficiency in English enhances individuals' capacity by providing access to numerous opportunities across several sectors. Besides, English functions as a lingua franca, facilitating effective communication, especially among individuals from diverse ethnicities with differing native languages. Beyond social and scholastic spheres, it is inconceivable to envision a professional identity devoid of English proficiency since globalization drives individuals to raise their linguistic competency (Kırmızı & Er, 2024). It is evident that English significantly influences individuals' lives, prompting them to reap the benefits during their lifespans.

A unique generative AI style, ChatGPT (Chat Generative Pre-Trained Transformer), emerged in 2022 through Open AI, demonstrating the ability to compose text closely resembling human language (Arefian et al., 2024; Setiawan & Rauf, 2025). It synthesises information from three main sources: readily available web materials, third-party contributions by partners, and users. It rose to fame so quickly that its users exceeded 100 million within two months of its launch (Kara & Kucuk, 2023; Pandey & Bhusal, 2024). Different versions of ChatGPT are 3, 3.5, 4, and 4o, with distinguishing qualities. For example, ChatGPT 3 is characterised by feeding a large amount of data and then fine-tuning it to perform specific tasks, such as translation, summarisation, or analysis. On the other hand, ChatGPT 3.5 has certain updates that are more secure, accountable, and interactive (Mohamed, 2024). In other words, this version welcomed human feedback to improve machine learning algorithms. In addition to the text-only mode, GPT 4 enhances graphics and natural language codes. Conversely, ChatGPT 4o has crystal clear differences from its predecessors. It is a brand-new model that can reason across video, audio, vision, and text in real time, which leads the business in transformative steps. With its enhanced processing speed, expanded multimedia support, and improved accessibility, ChatGPT 4o represents a significant step forward in developing language models (Kurudirek & Berdieva, 2024; Rezai et al., 2024). Its versatility and efficiency make it a valuable tool for apps, from content creation and customer support to research and analysis. It can be argued that ChatGPT has contributed exponentially to increasing the popularity of AI chatbots.

The notion of self-efficacy (SE) has been diligently explored to determine its role in learning. It is inspired by Bandura's (2008) Social Recognition Theory (SRT), which asserts that an individual's degree of trusting their abilities directly affects their performance. Simply put, a self-assured person can defy and overcome challenges, whereas low efficacy prevents people from taking action confidently, reducing the success rate sharply. Several factors play pivotal roles in shaping students' SE degrees. To name a few, past experiences can lower or raise confidence according to their implications (Genc et al., 2016; Kucuk, 2024). Previous success stories can boost motivation, while failures can hinder proactive behaviours. In addition, people around students or role models can affect students' confidence. In other words, acquaintances and influential people can shape your journey to be self-assured or shy. Next, verbal encouragement from peers, teachers, parents, or AI chatbots is irreplaceable, so its availability or deficiency directly impacts students' SE (Celik & Kara, 2024a; Namaziandost & Çakmak,

2020). Besides, the well-being of students plays an integral role in students' confidence because physical and psychological issues can potentially exacerbate the situation and decline SE. Considering the factors, it can be asserted that SE can be kept high by combining many elements; otherwise, it cannot be envisioned to notice the fruits of enhanced SE.

Developing speaking skills in a foreign language is one of the most enormous challenges that can affect a person's ability to initiate and conduct one-on-one conversations. In other words, speaking anxiety can become a hurdle that adversely affects students' performance. In this regard, meaningful interaction between speakers significantly promotes their speaking performance and reduces anxiety with well-established practices (Üstünbaş, 2024). Thus, learners may seize unparalleled opportunities to face challenges and devise sensible strategies in speaking circles. It is undeniable that speaking enhancement is impossible without making an incessant effort. Apart from physical speaking partners, GEN-AI tools, primarily ChatGPT, have gained increasing momentum to reap the advantages in many formats. For example, ChatGPT offers insightful points regarding pronunciation, coherence, relevancy, transition words, and logical sequence in an apprehension-free virtual atmosphere. Additionally, receiving feedback in a graded form based on students' levels is another milestone in the GEN-AI field. Additionally, recording conversations or texts provides instant feedback for learners to access invaluable information anytime, anywhere (Power, 2024). It is also observed that students exposed to constructive feedback by ChatGPT triple their chance to have more engaging talks with their peers in classes. More specifically, learners can consolidate their linguistic powers in oral tasks to strengthen their speaking SE with an easily understandable and manageable format. Thus, it is envisioned that speaking anxiety is reduced, and self-confidence is improved. All aforementioned features of ChatGPT distinguish it from other tools because it sets the standard for boosting the speaking SE of learners.

Literature Review

Studies on speaking SE have quadrupled in the 21st century because its effects on developing students' overall linguistic skills are acknowledged in many circles. Psychologist Albert Bandura pointed out in his highly acclaimed book *Self-efficacy: The Exercise of Control*, that learners' mental and emotional states should match each other harmoniously to get the maximum benefits. His input has contributed to the popularity of the SE concept in educational settings, igniting the curiosity of many researchers. Since then, several studies have focused on the role of parents (Albanese et al., 2019), peers (Bellò et al., 2018), teachers (Hajovsky et al., 2020), classroom atmosphere (Beri & Stanikzai, 2018), and virtual assistants (Wiboolyasarini & Jinowat, 2024; Wright, 2024) to gain adequate SE. For instance, Albanese et al. (2019) conducted a wide-scale study in the US, where they found that parents' well-being and supportive remarks positively affect students' overall SE. Another insightful study was carried out in Italy by Bellò et al. (2018), who uncovered that the way peers behave matters considerably in raising students' SE. Likewise, Hajovsky et al. (2020) stressed that teachers' SE is fundamental in instilling hope and motivation for American students. Similarly, Beri and Stanikzai (2018) underlined that a well-equipped classroom atmosphere urges the learners to exchange ideas and foster learning, substantially changing the SE level. Besides, Amin (2023) postulated that virtual assistants can transform education and offer endless potential at educational institutions based on his study in Iraq.

SE has taken increasing attention in the foreign language learning and teaching process. In this regard, integrating GEN-AI tools, primarily ChatGPT, into EFL classes becomes inevitable to grasp the opportunities in terms of macro- or micro-skills development, increased motivation, and enhanced language proficiency so the speaking SE levels of the learners can be sharpened

considerably. Besides, students' attitudes towards the language learning process through ChatGPT have been affected positively. The rationale for implementing AI-driven approaches in EFL classes is that chatbots assist learners in pinpointing their positive and negative sides while using more individualised conduct, monitoring students' development more closely, and offering more immediate and timely intervention (Jamoom & Bahron, 2024; Mulyono et al., 2019).

ChatGPT in Macro-Skills Development

GEN-AI tools have been studied thoroughly in EFL settings to measure their influence in the enhancement of macro-skills. To illustrate, Lin and Chen's study (2024) unearthed that ChatGPT can fulfil numerous functions in reading classes, from generating exams by considering the degree of difficulty to grading students' papers with constructive feedback in China at the tertiary level. His study revealed that composing reading exams in various formats considering students' levels captured educators' attention because they noticed positive effects on students' scores. Another notable study was conducted by Aryadoust et al. (2024), who found that ChatGPT can transform listening lessons by turning texts into audio tracks and generating various questions based on scripts. Their study showed that Singaporean high school students developed their listening skills substantially thanks to adopting AI-based instruction in listening sessions as they learned how to take notes and make interpretations based on the feedback. Subsequently, Syaripuddin's (2024) comprehensive study in Indonesia unearthed that language preparatory school students could boost oral proficiency considerably after using ChatGPT as a speaking tutor. Besides, Fitria's (2023) study in Indonesia on secondary school students documented that ChatGPT offers untapped potential in academic writing classes because it can be used to brainstorm, get model essays, change the degree of difficulty, get feedback about common mistakes, and customise texts based on the region.

ChatGPT in Micro-Skills Development

ChatGPT has been thoroughly investigated in the enhancement of micro-skills development. To name a few, El-Said (2024) explored the role of ChatGPT in elevating students' grammatical accuracy in Egypt at the tertiary level, which revealed that it sharpens learners' grammar by asking relevant questions and offering insightful comments to fix chronic grammar mistakes. In addition, Farr (2024) carried out a wide-scale study in the USA on students who took English for Specific Academic Purposes courses to measure the effectiveness of ChatGPT in empowering students' lexical powers. Her study showed that ChatGPT can improve students' vocabulary repertoire by creating vocabulary questions in varied formats, such as open-ended, fill-in-the-blank, and matching. After that, Kim and Park (2023) probed the impact of ChatGPT on enhancing students' pronunciation skills in Korea, where students could master pronunciation thanks to comprehensible and motivating inputs. It can be argued that ChatGPT is promising to raise students' capacity in terms of micro-skills.

ChatGPT in Increasing Overall Motivation and Language Proficiency

ChatGPT's role in enhancing motivation and increasing the language proficiency level of EFL learners has been studied extensively by numerous researchers around the world. For instance, Ali et al. (2023) have figured out that ChatGPT has unique features to trigger students' intrinsic and extrinsic motivation, thereby boosting Saudi students' overall language proficiency in higher education because it employs a holistic approach to sharpen learners' competence gradually. Another noteworthy study was carried out by Balcı (2024), whose study in Turkey at the tertiary level uncovered that learners and pre-service teachers take advantage of ChatGPT

distinctively because individualised feedback, balanced sequence of knowledge, and variety increases students' enthusiasm and overall proficiency in the long run. Likewise, Yamaoka's (2024) insightful study showed that Japanese university students reaped the benefits of ChatGPT as a motivational source; therefore, they upgraded their language proficiency.

ChatGPT as a Speaking Tutor

With the advent of ChatGPT's voice-oriented prompts in the conversation mode, its usage has expanded to train learners to be more self-assured in their speaking attempts. In this regard, a growing number of publications have been released by renowned authors. To illustrate, Limo et al. (2023) investigated the influence of ChatGPT in raising their speaking performance as a virtual tutor in Colombia on first-year university students. The findings revealed that using the tool as a complementary speaking tool helped students reduce anxiety and speak more fluently. Additionally, Jeon et al. (2023) stated that receiving feedback through a speech recognition system multiplied the value of AI chatbots, particularly ChatGPT because they support students' learning incessantly with science-driven and pedagogically approved strategies. To substantiate other studies, Hellmich et al. (2024) conducted a comprehensive study in a US university where respondents enrolled in 20 language learning courses, the results of which documented that ChatGPT has certain advantages over others due to its instant reply-encompassing creativity, emotional intelligence, and accuracy.

ChatGPT-Oriented Studies in Iraq

An increasing number of studies have been launched and finalised to measure the effects of ChatGPT on developing EFL learners' numerous skills in Iraq where the present study was carried out. Researchers have conducted studies to contribute to existing literature with practical implementations. To name a few, Kara (2024) carried out a multi-disciplinary study on first-year students whose lessons were enriched with ecology-oriented news accompanied by reading tests created by ChatGPT. The findings showed that ChatGPT has far-reaching effects in increasing students' reading scores with catchy news and relevant questions based on the news script. Subsequently, Amin's (2023) study hinted that ChatGPT facilitates personalised language learning, so it has untapped potential in increasing students' overall development in EFL settings. Additionally, Bilgin and Yildiz's (2024) study showed that AI tools, particularly ChatGPT, can enhance students' 21st-century skills at all stages of the university, so they can improve their English faster than their peers when they learn how to undertake responsibility and initiative through AI. Another notable study was conducted by Kucuk (2023), whose findings documented that ChatGPT paves the way for sharpening grammatical accuracy at the tertiary level thanks to its instant and accurate responses to any queries. Besides, Aldahhan and Razak's (2024) study revealed that speaking classes should be enriched with AI tools to foster students' learning in a welcoming atmosphere. The studies mentioned above display that AI tools have been implemented in increasing numbers at educational institutions in Iraq; however, the current study will be the first attempt to measure the influence of ChatGPT in elevating Iraqi students' speaking SE, thereby speaking more fluently at the tertiary level. In this regard, the following research questions given below served as the basis for this study:

RQ1: How influential is ChatGPT in developing the speaking SE of learners?

RQ2: Does ChatGPT positively affect students' speaking scores as a tutor?

RQ3: What are the perceptions of learners about ChatGPT enriched speaking instruction?

Method

Research Design

The study's primary objective is to investigate the impact of ChatGPT on developing speaking SE of learners of English as a Foreign Language. To attain this goal, a mixed-methods research design was employed. Data collection encompassed both quantitative and qualitative methodologies via a sequential explanatory model consisting of two separate phases: a qualitative phase was preceded by a quantitative period (Song et al., 2024; Timans et al., 2019). This strategy is employed with a common goal of gaining numerical and verbal data, which can be triangulated to determine whether they match each other harmoniously regarding the research question. Thus, it is not unsurprising that more reliable interpretations can be uncovered (Bowen et al., 2017).

Setting and Participants

The population of the study is a renowned private university in Erbil, Iraq, where students from various nationalities have received top-notch education from qualified lecturers through English Medium Instruction since 2008. Students can take numerous courses to sharpen their English skills, thereby being qualified to be employed with a satisfactory income upon graduation. They can participate in a dynamic and professional cycle at the preparatory school where they systematically develop all macro- and micro-skills with a well-established and monitored curriculum. Besides, they can take Foundation English courses to master English and increase their chance of obtaining a high GPA in their departments because the foundation year matters to grasping all details in departmental courses and undertaking responsibility through various assignments, such as reports, presentations, field trips, or essays. Subsequently, students enroll in Advanced English courses to pass the thresholds and become proficient in English as they can figure out their weaknesses in English and be granted the opportunity to fix all issues in the way of earning proficiency. Considering those facts, the participants of the study were chosen from sophomore students who took Advanced English courses in the 2024–2025 academic year. Their courses were covered by a renowned series of Cambridge University Press 'OWN IT.' This series was first released in 2019 and instantly captured the EFL world's attention because it is a complete set with videos, games, worksheets, and I-tools. It is worth mentioning that the books also have assessment tools to measure students' learning after each unit, semester, and academic year, respectively, in paper-based and online formats. Additionally, the books were designed to draw the attention of adults with engaging topics about social media, technology, and AI, so students had many common hobbies with the characters in the books. Those advantages drove the stakeholders to choose the book for the Advanced English course. Participants ($N = 44$) were chosen from the population through convenience sampling thanks to its nature to select participants based on convenience. Both control and experimental group students were studying in the pharmacy department. Their ages ranged from 18 to 29. The number of female students ($n = 18$) was less than male students ($n = 26$).

The rationale for choosing Advanced English students is that their proficiency levels were adequate to easily adapt to ChatGPT-enriched speaking sessions. Additionally, they were familiar with ICT tools because they had previously taken several courses. After that, they had background knowledge in English through the Foundation English course for the whole year. Thus, they did not have any technical or academic issues to face challenges and deal with wisely. In light of this, the researchers invited the participants to participate voluntarily. They signed the consent forms, took the workshops, and joined the preliminary Cambridge Proficiency Test

(CPT) to place them homogeneously. Having examined the results, it was observed that their levels were similar to each other in the prior phase of the treatment period. In other words, their levels were recorded as B1 based on the Common European Framework of Reference for Languages (CEFR).

Procedure

Control group students' lessons were covered based on the official curriculum; however, experimental group students' lessons were enriched with speaking sessions. In other words, control group students took part in lessons and completed all activities in a graded way, ranging from expressing their ideas to taking relevant tests. However, experimental group students' lessons were uncommon because they practised at home with ChatGPT as the speaking tutor and did more speaking exercises in class with the lecturer. They underwent several stages to implement the instruction flawlessly and reap the benefits correspondingly. To illustrate, students scanned the book thoroughly with the lecturer to choose worthy themes between units 1 and 3. Later, they sequenced the topics to cover week by week. Subsequently, students were trained to use ChatGPT as the speaking tutor. They also received intensive training about downloading the app, creating the account, logging in several times, and activating distinguishing features with appropriate prompts. The lecturer also rehearsed ChatGPT to be a role model and eliminate possible hindrances in advance. Subsequently, students inserted the questions into ChatGPT, which directed some of them to students weekly. When students shot their responses, ChatGPT offered valuable feedback about delivery, language use, topic development, and pronunciation. More specifically, ChatGPT listened to students' responses, found some mistakes, informed them constructively, and offered suggestions for improvement. This cycle continued for 8 weeks. At the end of the cycle, students chose one example and submitted it to the lecturer. The lecturer played them in the class to check the effectiveness of the instruction mode.

Speaking SE surveys and speaking exams were integrated into the study to collect data and make interpretations accordingly. SE surveys were created based on Zhang et al.'s (2023) scale to get learners' perspectives on speaking SE. Their scale fills the void in the EFL domain, where speaking SE tools are scarce. The speaking SE scale encompassed four dimensions: planning the conversation, shooting the responses, language structure, and listeners with 16 items in a 5-point Likert scale mode. Their speaking SE scores were compared to determine whether the intervention period enriched by ChatGPT as a speaking tutor affected their speaking SE degrees. The subsequent data collection tools were speaking exams, which EFL experts created to gauge students' speaking levels in pre-defined intervals. Both students' speaking exam scores and surveys were analysed using SPSS 29. In the final moment, the lecturer received students' personal opinions about the whole study with focus group interviews, some of which were unleashed in the findings section to have a clear image.

The interview was split into certain stages to ensure professional conduct. Firstly, the researchers composed the questions. After that, a vetting process was applied with the help of six EFL experts knowledgeable about linguistics and statistical analysis. Interview questions were refined according to their comments, so possible barriers against understanding were removed. After the initial screening, the interview was finalised. After that, interview questions were tested on other students as a pilot study, so possible issues were determined in advance. These procedures helped the researcher to ensure efficiency and excellence.

During the treatment period, students expressed their ideas freely within their classes. Each student had a say in their reserved time to avoid missing any valuable ideas. A 15-minute session was allocated for each student to shoot their responses independently. Additionally, all items were in English because their levels were quite satisfactory to understand all instructions

Table 1 *Weekly speaking topics*

Topic	Weeks
Describe your favorite social media platform	1
Describe ways to curb internet addiction	2
Describe a typical day in your life	3
Depict your favorite app	4
Illustrate the best friend with physical and personal qualities	5
Visualise moments in childhood	6
Explore the best movie	7
Elaborate favorite music genre with specific examples	8



Figure 1 *Stages of conducting a ChatGPT enriched speaking instruction.*

in English. Accordingly, students’ ideas were recorded, transcribed, refined, categorised into meaningful segments, and coded, which aligned with the principles of content analysis. Later, it was time to hold speaking exams to gain quantitative data so researchers could gather meaningful data. Speaking topics which were explored throughout the study were given below:

As illustrated in Table 1, speaking topics were relevant to their real lives, so it was expected to raise their motivation to speak more confidently. The sequence of the study is also illustrated in Figure 1 below:

Validity, Reliability, and Ethics

The researcher employed several strategies to ensure validity, reliability, and ethical excellence. For instance, the speaking SE scale and speaking exams were piloted on other students whose characteristics were not different from the participants. This endeavour was undertaken to

Table 2 *Piloting scales and exams through cronbach alpha*

Cronbach Alpha Index	
Speaking self-efficacy questionnaire	.93
Speaking exams	.89

guarantee internal consistency. In the next phase, the face and content validity of the items were evaluated by experts to refine and increase the intelligibility of scale and exam. According to experts’ insightful comments, items were revised and modified correspondingly. Having completed these tasks, the researcher informed the participants to get their consent once more as they were notified that they were able to withdraw from the study at any time without making a credible justification. Additionally, the university research committee examined the study several times and approved the request to start and finalise the process. The results of the Cronbach Alpha Index are illustrated in Table 2 below:

Table 2 shows that the SE scale and speaking exam met the reliability criteria with .93 and .89, respectively. According to Forero (2023), Cronbach Alpha should be .70 and above to be considered reliable, which was much higher in this study. Thus, the researcher was inclined to proceed to the next stage.

Findings

Findings were elaborated under two headings: quantitative and qualitative data. The former was formed through questionnaires and speaking exams, whereas the latter was compiled from students’ personal opinions in the focus group interviews.

Quantitative Data

Quantitative data were collected to address the first research question, which was designed to measure the differences in speaking more confidently via ChatGPT-enriched instruction. Correspondingly, speaking SE questionnaires and exams were held. Table 3 illustrates the details of the questionnaire.

The SE of the learners was investigated through the questionnaires at two intervals. When pre-test scores were probed, it was noticed that all participants had issues with all aspects of speaking regardless of their groups. Sub-categories of speaking SE were unearthed, and they had severe

Table 3 *Participants’ level of speaking self-efficacy in english*

Aspects of speaking self-efficacy	Before intervention				After intervention			
	Control Mean SD		Experimental Mean SD		Control Mean SD		Experimental Mean SD	
Planning	2.91	1.34	3.00	1.34	3.14	1.32	4.32	.839
Shooting	2.77	1.37	2.73	1.27	2.82	1.46	4.41	.959
Language Structure	2.55	1.22	2.59	1.14	2.77	1.34	4.36	1.04
Listeners	2.73	1.03	2.64	1.29	3.05	1.43	4.09	.750
Overall	2.74	1.23	2.91	1.31	2.94	1.37	4.30	.897

Table 4 *The analysis of pre-test and post-test speaking exam scores*

Variables	Groups	N	Mean	SD	t	df	p
Pre-test	Control	22	53.41	14.75			
Pre-test	Experimental	22	54.32	14.16	.208	42	.836
Post-test	Control	22	57.05	16.15			
Post-test	Experimental	22	80.00	10.69	.557	42	.000

problems regarding fluency, coherence, pronunciation, grammatical range, and lexical resources. Their mean scores varied from 2.55 to 3.00, substantiating their low SE beliefs. On the other hand, using ChatGPT as a speaking tutor paved the way for experimental group students. They increased their mean scores substantially in all categories. To name a few, their mean scores leapt from 3.00 to 4.32 in the planning category. Additionally, students improved their performance in addressing their speech section, considering the mean scores from 2.73 to 4.41. Subsequently, using ChatGPT as a speaking tutor developed students' performance in fixing linguistic errors. Thus, their mean scores elevated from 2.59 to 4.36. After that, they learned how to control their speaking anxiety towards the audience. Their mean scores rose from 2.64 to 4.09, which was significant. Once all items were analysed, it was clear that using ChatGPT as the speaking tutor sharpened students' SE beliefs, thereby increasing their chance to eliminate barriers and speak more confidently. Overall mean scores increased from 2.91 to 4.30. However, control group students' SE scores were far from being noticeable. In other words, covering speaking classes in a traditional mode was not as efficient as technology-enriched speaking sessions.

Apart from the questionnaires, speaking exams were integrated into the study to gauge the effects of the treatment period through statistical means. Students responded to the speaking queries in intervals. It is worth mentioning that speaking questions were chosen from the pool created by the team based on common topics of the coursebook, OWN IT 3. Three EFL experts, along with the researchers, joined the exams to grade performance based on the rubric. Subsequently, the exam committee considered the grades with respective notes and decided on the final points. The statistical analysis of the exam is depicted in Table 4 in the mode of pre-test and post-test.

The numerical data in Table 4 show that students' levels were approximately the same in pre-test scores. Mean scores were roughly the same, with 53.41 for the control group and 54.32 for the experimental group. In other words, the *p*-value of .836 is higher than 0.05, which wasn't significant. On the other hand, a considerable leap was measured in post-test scores considering the *p*-value of .000, which is less than 0.05. In other words, control group students increased their scores by only 4 points, while it was 27 points in the experimental group thanks to adopting the AI-enriched speaking instruction.

Qualitative Data

Questionnaires and exam results were triangulated with semi-structured interviews to determine whether findings matched each other harmoniously without leaving room for hesitation. In this regard, the study's effectiveness was investigated thoroughly. Participants responded to the given questions:

- What is your personal opinion about improving speaking SE through ChatGPT?
- How did this study affect your perspective toward sharpening speaking SE via ChatGPT and earning high grades in speaking exams?

Table 5 Themes to reflect learners’ opinions on using ChatGPT as a tutor

Themes	f	%
Speaking more confidently	20	90
Developing speaking proficiency	21	95
Deriving pleasure	19	86
Flexibility	22	100
Insufficient responses on feeling	9	40

Participants’ opinions were saved under five themes: speaking assertively, improving overall proficiency, getting pleasure, using conveniently, and suffering from a lack of emotional intelligence. Table 5 depicts the meticulous details below:

Once Table 5 is examined in detail, it is crystal clear that using ChatGPT as a speaking tutor received primarily positive feedback except for the emotional intelligence dimension. For instance, most respondents (90%) reiterated that they started speaking more assertively after being exposed to constructive feedback by ChatGPT.

(5) *“I had speaking anxiety in previous classes because I thought that my friends would humiliate or make fun of me when I made any mistakes. I handled this fear concurrently after using ChatGPT as a speaking tutor. I observed that speaking was not as tough as I expected. Additionally, getting recommendations to polish my English increased my motivation to communicate more assertively. Overcoming speaking barriers and fixing minor issues were my gains in this study.”*

Additionally, nearly all respondents (95%) postulated that using ChatGPT improved their proficiency level.

(6) *“I could not trust my speaking proficiency in the past because I was inclined to make a mistake and pause suddenly. Using ChatGPT as a speaking tutor has helped me eliminate such issues. The more I practised with it, the higher the proficiency level I reached. Meanwhile, I transferred my speaking knowledge to other skills such as reading, writing, and listening because the words I learned were essential to conducting conversations in speaking, diving into reading passages, comprehending listening tracks easily, and expanding my reasons of support with clear examples in writing. Overall, this study positively affected my overall proficiency.”*

Besides, getting pleasure (86%) prompted the respondents to join the sessions willingly because it was an unprecedented experience for them to embark on a journey where they improved their speaking skills incessantly.

(9) *“I do not like monotony, so I would like to apply this to my life, including education. Accordingly, ChatGPT was a real saver, allowing me to converse and get novel responses. ChatGPT could balance the degree of difficulty in offering feedback and shooting model responses, thereby increasing my satisfaction rate. I progressed, considering my current level and capabilities. After that, paying attention to my intonation, pace, sentence structure, and coherence drove me to speak accurately, so I derived so much pleasure while being tutored by ChatGPT.”*

Subsequently, being privileged to take advantage of ChatGPT flexibly earned the respect of all respondents (100%).

(12) *“I am a big fan of hybrid education, so I want to get educated at schools and online simultaneously. In this regard, ChatGPT fits my specific needs. I learned some concepts in*

my regular classes but wanted to go further through ChatGPT. We studied the topics with my dedicated lecturer, and I stimulated my creativity via ChatGPT because it offered insightful comments to me tirelessly. Furthermore, I could carry out conversations at any time, so it fostered my learning. I am into learning flexibly, which was put into practice by ChatGPT. What triggered my attention in this study was that I could rewind and listen to the recordings as much as possible, so I learned another tip as I played the recording again."

The fifth theme was the weakness of the ChatGPT in imitating the human voice as a trained chatbot, thereby having certain problems passing on emotion. Thus, several students (40%) stated that there was a vast emotional intelligence gap between the actual lecturers and ChatGPT, so they could not always maintain their concentration. However, they also praised the improvement of ChatGPT in closing the gaps between human voice and AI chatbots as much as possible.

(14) "I thought that chatbots could not be replaced with any real teacher. On the other hand, I am not so strict about my perspective now because ChatGPT has improved since its inception. Its emotional enhancement motivated me to be more engaged. It created a welcoming atmosphere, offered constructive feedback, and directed me to certain websites to improve my pronunciation and overall delivery. Using ChatGPT as a speaking tutor can also positively change the classroom atmosphere because we have the background to compose our ideas in conversations and shoot our responses confidently. I mean that ChatGPT can increase its efficiency in the eyes of the users when the feelings of the users are observed and responses are shaped accordingly."

Discussion

This study, which used a mixed-methods research design, sought to determine ChatGPT's effectiveness in enhancing the speaking SE of EFL learners, thereby increasing their speaking scores, so research questions were designed and addressed accordingly.

The first research question was addressed through speaking exams, which showed that ChatGPT was a marvelous guide to developing speaking skills. More specifically, reducing speaking anxiety, using grammar accurately, pronouncing the words precisely, and using a wide range of vocabulary to deliver the expressions coherently ensured considerable rise. Therefore, the learners' SE was positively affected by their having more oral proficiency. Those findings dovetail with the studies of Farr (2024) and Kara (2024), which documented that ChatGPT teaches users how to manage time, overcome stress, enrich their expressions, and fix language mistakes, so it is not uncommon that they speak more assertively with a persuasive tone. Similarly, Bilgin and Yildiz (2024) pointed out that AI tools help learners sharpen their 21st-century skills, enhancing their leadership and problem-solving skills, thereby enhancing speaking more confidently and enthusiastically. To substantiate other studies, Sitepu et al. (2024) stated that AI tools should not be seen as a foe. In contrast, they should be recognised as a companion to support learning with unprecedented and proven strategies.

The second research question was examined through questionnaires and interviews to receive students' genuine opinions. In this regard, questionnaires documented that using ChatGPT as a speaking tutor paved the way for speaking more confidently, accurately, and convincingly, thanks to the nature of ChatGPT to offer constructive feedback. This finding resonates with the studies of Abdelaal and Al Sawy (2024) and Escarlos and Tan (2017). They asserted that AI chatbots, primarily ChatGPT, have enormous potential to increase students' speaking SE with guided and graded exercises because they offer individualised feedback in an engaging, stress-free environment. Likewise, El-Said (2024) stated that new generations are digital natives, so they derive much pleasure while delving into technological tools in communicative acts

through ChatGPT. Their motivation rises when they find answers to their queries in an easily understandable and retrievable format. Complementary findings in the interviews aligned with the questionnaires, which displayed deriving pleasure during meaningful conversations and having the freedom to learn flexibly at any time, were articulated as the distinguishing qualities of the ChatGPT-enriched speaking instruction. This finding aligned with Karam's (2024) study, which unearthed that AI tools, specifically ChatGPT, can increase learners' satisfaction rates with novel and relevant responses without being restricted to time and setting.

A negative theme arose in the interview about the lack of emotion in reactions. This is consistent with the study of Fitria (2023), which hinted that actual human interaction cannot be envisioned with ChatGPT despite being trained to imitate human conversations as much as possible. Aligned with other studies, Celik and Kara (2024b) asserted that ChatGPT should not assume responsibility in educational settings because real teachers cannot be underestimated with their distinctive and irreplaceable roles. When all findings were triangulated, it was also observed that ChatGPT met participants' needs precisely, so their speaking SE was boosted correspondingly.

Conclusion

This research purported to explore the potential effects of ChatGPT as a speaking tutor in elevating SE at the tertiary level in Iraq through speaking exams, questionnaires, and interviews. All findings uncovered that ChatGPT-enriched speaking instruction is not a threat to a physical teacher; conversely, it has numerous benefits in teaching English engagingly thanks to its ability to improve students' speaking skills with balanced, coherent, relevant conversations in a flexible mode. Additionally, breaking monotony with novel approaches increased students' enthusiasm to get pleasure and explore further possibilities. Another benefit of the treatment was that students handled speaking anxiety and spoke more confidently with convincing remarks agreed upon by all participants of the experimental group in the interview unanimously. Adopting a communicative language teaching method, using technology harmoniously, assigning extracurricular tasks, and offering feedback based on students' recordings with ChatGPT paved the way for overcoming speaking SE issues and articulating ideas more self-assuredly. On the other hand, covering speaking activities solely in class by the lecturer did not yield better results. They could not improve their speaking scores significantly, which was also substantiated in the complementary surveys.

Pedagogical Implications and Suggestions for Future Studies

The results of the study have several implications for educational institutions. Artificial intelligence tools can be integrated into EFL classes as a practical method to enhance students' speaking SE. This attempt can pave the way for noticing drastic changes in their communicative competence. To realise this successfully in EFL classes, intensive in-service training sessions should be held with practical implementations to see possible problems and fix them wisely. In this regard, equipping the courses with the required innovative technologies and installing programs in advance can increase the productivity of the students dramatically. Besides, assessment criteria should be adjusted to leave room for online assignments, including using ChatGPT as a speaking tutor so that students can be encouraged to join online sessions and grab opportunities to sharpen their English. Afterwards, parent meetings can be held to inform them about the advantages of ChatGPT-integrated EFL classes and show practical strategies to help their kids carry out their speaking assignments appropriately. Finally, students' recordings can be shared with policymakers to see the untapped potential of AI tools in education and take measures to invest more in having technology-enriched classes as standard rather than being scattered in

only certain regions. All these attempts can tailor students' specific needs and increase their first-hand experience of learning English enjoyably in Iraq and globally. Thus, speaking SE can be enhanced substantially, eliminating many forms of language learning anxiety.

The current study suffered from various limitations that reduced the findings' generalisability. Firstly, this study was limited to sophomore students at a private university in Erbil, Iraq, which can be expanded with other academic stages and institutions in the country. Besides, this study attempted to examine the role of ChatGPT in boosting speaking SE, which can be enlarged with other primary skills because it is a multi-faceted chatbot that meets learners' needs in many skills. Further studies can be conducted to examine whether age, gender, and level of learners significantly affect their speaking scores. Subsequently, this study only reflected the sincere opinions of learners, which can be extended to teachers who ran ChatGPT-enriched education. Finally, this study welcomed ChatGPT as an e-tutor, which can be compared with other AI tools like Gemini, Bard, and Perplexity. These new lines of inquiry will contribute to existing studies with different dimensions and capabilities.

References

- Abdelaal, N., & Al Sawy, I. (2024). Perceptions, challenges, and prospects: University professors' use of artificial intelligence in education. *Australian Journal of Applied Linguistics*, 7(1), 1309. <https://doi.org/10.29140/ajal.v7n1.1309>
- Albanese, A. M., Russo, G. R., & Geller, P. A. (2019). The role of parental self-efficacy in parent and child well-being: A systematic review of associated outcomes. *Child: Care, Health and Development*, 45(3), 333–363. <https://doi.org/10.1111/cch.12661>
- Aldahhan, A., & Razak, N. A. (2024). The impact of digital platforms on students' communicative competence of Iraqi undergraduate EFL students. *Migration Letters*, 21(7), 1007–1015.
- Ali, J. K. M., Shamsan, M. A. A., Hezam, T. A., & Mohammed, A. A. Q. (2023). Impact of ChatGPT on learning motivation: Teachers and students' voices. *Journal of English Studies in Arabia Felix*, 2(1), 41–49. <https://doi.org/10.56540/jesaf.v2i1.51>
- Amin, M. Y. M. (2023). AI and Chat GPT in language teaching: Enhancing EFL classroom support and transforming assessment techniques. *International Journal of Higher Education Pedagogies*, 4(4), 1–15. <https://doi.org/10.33422/ijhep.v4i4.554>
- Arefian, M. H., Çomoğlu, I., & Dikilitaş, K. (2024). Understanding EFL teachers' experiences of ChatGPT-driven collaborative reflective practice through a community of practice lens. *Innovation in Language Learning and Teaching*, 1–16. <https://doi.org/10.1080/17501229.2024.2412769>
- Aryadoust, V., Zakaria, A., & Jia, Y. (2024). Investigating the affordances of OpenAI's large language model in developing listening assessments. *Computers and Education: Artificial Intelligence*, 6, 100204. <https://doi.org/10.1016/j.caeai.2024.100204>
- Balci, Ö. (2024). The role of ChatGPT in English as a Foreign Language (EFL) learning and teaching: A systematic review. *International Journal of Current Educational Studies*, 3(1), 1–15. <https://doi.org/10.5281/zenodo.12544675>
- Bandura, A. (2008). The reconstrual of “free will” from the agentic perspective of social cognitive theory. *Are We Free*, 34(8), 86–127.
- Bellò, B., Mattana, V., & Loi, M. (2018). The power of peers: A new look at the impact of creativity, social context and self-efficacy on entrepreneurial intentions. *International Journal of Entrepreneurial Behavior & Research*, 24(1), 214–233. <https://doi.org/10.1108/ijeb-07-2016-0205>
- Beri, N., & Stanikzai, M. I. (2018). Self-efficacy beliefs, student engagement and learning in the classroom: A review paper. *American International Journal of Research in Humanities, Arts and Social Sciences*, 22(1), 213–222.
- Bilgin, R., & Yildiz, Y. (2024). Fostering EFL learners' writing competence through problem-based learning. *Forum for Linguistic Studies*, 6(3), 574–591. <https://doi.org/10.30564/fls.v6i3.6824>
- Bowen, P. W., Rose, R., & Pilkington, A. (2017). Mixed methods-theory and practice. Sequential, explanatory approach. *International Journal of Quantitative and Qualitative Research Methods*, 5(2), 10–25.

- Celik, B., & Kara, S. (2024a). Perceptions of parents on Raz-kid's role in enhancing students' micro and macro-skills holistically: a study on Ishik Brayaty International Primary School in Erbil. *Forum for Linguistic Studies*, 6(3), 558–573. <https://doi.org/10.30564/fls.v6i3.6822>
- Celik, B., & Kara, S. (2024b). Reaping the fruits of technology-integrated grammar instruction in EFL classes at the tertiary level through web 2.0 tools. *Novitas-ROYAL (Research on Youth and Language)*, 18(1), 18–36. <https://doi.org/10.5281/zenodo.10973103>
- El-Said, M. A. (2024). Using ChatGPT to enhance the third preparatory students' performance in English grammar. *CDELT Occasional Papers in the Development of English Education*, 86(1), 291–316. <https://doi.org/10.21608/opde.2024.362826>
- Escarros, G. S., & Tan, D. (2017). Motives, attitudes and performance of teacher education students. *International Journal of Scientific & Technology Research*, 6(10), 20–25.
- Farr, C. (2024). *Unmasking ChatGPT: The Challenges of using artificial intelligence for learning vocabulary in English as an additional language* (Doctoral dissertation, Canada). <http://hdl.handle.net/1828/15835>
- Fitria, T. N. (2023). Artificial intelligence (AI) technology in OpenAI ChatGPT application: A review of ChatGPT in writing English essay. *ELT Forum: Journal of English Language Teaching*, 12(1), 44–58. <https://doi.org/10.15294/elt.v12i1.64069>
- Forero, C. G. (2023). Cronbach's Alpha. In: Maggino, F. (eds) *Encyclopedia of Quality of Life and Well-Being Research*. Springer, Cham. https://doi.org/10.1007/978-3-031-17299-1_622
- Genc, G., Kulusakli, E., & Aydin, S. (2016). Exploring EFL learners' perceived self-efficacy and beliefs on English language learning. *Australian Journal of Teacher Education*, 41(2), 53–68. <https://search.informit.org/doi/abs/10.3316/informit.855779233896679>
- Hajovsky, D. B., Chesnut, S. R., & Jensen, K. M. (2020). The role of teachers' self-efficacy beliefs in the development of teacher-student relationships. *Journal of School Psychology*, 82, 141–158. <https://doi.org/10.1016/j.jsp.2020.09.001>
- Hellmich, E. A., Vinall, K., Brandt, Z. M., Chen, S., & Sparks, M. M. (2024). ChatGPT in language education: Centering learner voices. *Technology in Language Teaching & Learning*, 6(3), 1741. <https://doi.org/10.29140/tltl.v6n3.1741>
- Issa, F. M. (2024). Using infographics as an educational technology tool in EFL writing: University of Baghdad case study. *Arab World English Journal*, 15(1), 166–181. <https://dx.doi.org/10.24093/awej/vol15no1.11>
- Jamoom, O. A., & Bahron, N. O. (2024). From hesitation to fluency: Unraveling speaking barriers in EFL context. *International Journal of English Literature and Social Sciences*, 9(5), 102–110. <https://ijels.journallit.com/index.php/ijels/article/view/234>
- Jeon, J., Lee, S., & Choe, H. (2023). Beyond ChatGPT: A conceptual framework and systematic review of speech-recognition chatbots for language learning. *Computers & Education*, 104898. <https://doi.org/10.1016/j.compedu.2023.104898>
- Kara, S., & Kucuk, T. (2023). The effects of picture dictionaries in promoting vocabulary development of EFL learners at tertiary level. *Novitas-ROYAL (Research on Youth and Language)*, 17(2), 80–94. <https://doi.org/10.5281/zenodo.10015826>
- Kara, S. (2024). Enhancing students' writing competence and 21st-century skills simultaneously via infographics. *Australian Journal of Applied Linguistics*, 7(3), 2044–2044. <https://doi.org/10.29140/ajal.v7n3.2044>
- Karam, Z. N. (2024). First-year students' attitude towards English as a medium of instruction: a study on Tishk International University, Erbil, Iraq. *Arab World English Journal*, 15(4), 189–204. <https://dx.doi.org/10.24093/awej/vol15no4.12>
- Kirmizi, Ö., & Er, H. K. (2024). An examination of the interaction patterns and argumentative quality in collaborative writing. *Novitas-ROYAL (Research on Youth and Language)*, 18(2), 254–269. <https://doi.org/10.5281/zenodo.13924255>
- Kim, S., & Park, S. H. (2023). Young Korean EFL learners' perception of role-playing scripts: ChatGPT vs. textbooks. *Korean Journal of English Language and Linguistics*, 23(2), 1136–1153.
- Kucuk, T. (2023). Students' perceptions of the use of ICT tools in language preparatory school. *Arab World English Journal*, 14(4), 197–212. <https://dx.doi.org/10.24093/awej/vol14no4.12>

- Kucuk, T. (2024). ChatGPT integrated grammar teaching and learning in EFL classes: a study on the international university students in Erbil, Iraq. *Arab World English Journal (AWEJ)* Special Issue on ChatGPT, April 2024: 100–111. <https://dx.doi.org/10.24093/aweJ/ChatGPT.6>
- Kurudirek, A., & Berdieva, O. (2024). The more empowered schools, the more fruitful students. *Educenter: Jurnal Ilmiah Pendidikan*, 3(1), 1–9. <https://doi.org/10.55904/educenter.v3i1.921>
- Limo, F. A. F., Tiza, D. R. H., Roque, M. M., Herrera, E. E., Murillo, J. P. M., Huallpa, & Gonzáles, J. L. A. (2023). Personalized tutoring: ChatGPT as a virtual tutor for personalized learning experiences. *Przestrzeń Społeczna (Social Space)*, 23(1), 293–312. <https://socialspacejournal.eu/menu-script/index.php/ssj/article/view/176>
- Lin, Z., & Chen, H. (2024). Investigating the capability of ChatGPT for generating multiple-choice reading comprehension items. *System*, 123, 103344. <https://doi.org/10.1016/j.system.2024.103344>
- Mohamed, A. M. (2024). Exploring the potential of an AI-based Chatbot (ChatGPT) in enhancing English as a Foreign Language (EFL) teaching: perceptions of EFL faculty members. *Education and Information Technologies*, 29(3), 3195–3217. <https://doi.org/10.1007/s10639-023-11917-z>
- Mulyono, H., Sari, R. N., & Ningsih, S. K. (2019). Factors contributing to EFL students' speaking anxiety. *Register Journal*, 12(1), 13–27. <http://dx.doi.org/10.18326/rgt.v12i1.13-27>
- Namaziandost, E., & Çakmak, F. (2020). An account of EFL learners' self-efficacy and gender in the flipped classroom model. *Education and Information Technologies*, 25(5), 4041–4055. <https://doi.org/10.1007/s10639-020-10167-7>
- Pandey, H. L., & Bhusal, P. C. (2024). ChatGPT literacy for fostering language proficiency and writing skills in ESL/EFL classrooms. *Nepal Journal of Multidisciplinary Research*, 7(3), 1–24. <https://doi.org/10.3126/njmr.v7i3.70859>
- Power, R. (2024). Evaluating graduate education students' self-efficacy with the use of artificial intelligence agents: A case study. *Journal of Educational Informatics*, 5(1), 3–19. <https://journalofeducationalinformatics.ca/index.php/JEI/article/view/269>
- Rezai, A., Namaziandost, E., & Hwang, G. J. (2024). How can ChatGPT open promising avenues for L2 development? A phenomenological study involving EFL university students in Iran. *Computers in Human Behavior Reports*, 100510. <https://doi.org/10.1016/j.chbr.2024.100510>
- Setiawan, T., & Rauf, W. (2025). Google Assistant implementation for listening classes in senior high school in Barru: An exploratory case study. *Jurnal Edukasi Saintifik*, 5(1), 27–37.
- Sitepu, J., Lumbanraja, P., Silalahi, A. S., & Siahaan, E. (2024). Teaching digital service skills and innovative behavior to fresh graduates: An empirical analysis. *Novitas-ROYAL (Research on Youth and Language)*, 18(2), 220–229. <https://doi.org/10.5281/zenodo.13894294>
- Song, Y., Kim, J., Liu, Z., Li, C., & Xing, W. (2024). Students' perceived roles, opportunities, and challenges of a generative AI-powered teachable agent: a case of middle school math class. *Journal of Research on Technology in Education*, 1–19. <https://doi.org/10.1080/15391523.2024.2447727>
- Syaripuddin, R. (2024). Enhancing EFL students' oral proficiency: The ChatGPT-assisted debate clinic method. *JELITA*, 5(2), 627–639. <https://doi.org/10.56185/jelita.v5i2.651>
- Timans, R., Wouters, P., & Heilbron, J. (2019). Mixed methods research: what it is and what it could be. *Theory and Society*, 48, 193–216. <https://doi.org/10.1007/s11186-019-09345-5>
- Üstünbaş, Ü. (2024). Hey, GPT, can we have a chat?: A case study on EFL learners' AI speaking practice. *International Journal of Modern Education Studies*, 8(1), 91–107. <https://doi.org/10.51383/ijonmes.2024.318>
- Wiboolyasarini, W., & Jinowat, N. (2024). Exploring teachers' experiences in bilingual education for young learners: Implications for dual-language learning apps design. *Iranian Journal of Language Teaching Research*, 12(1), 45–64. <https://doi.org/10.30466/ijltr.2024.121417>
- Wright, N. (2024). Towards becoming a transnational language educator. *Australian Journal of Applied Linguistics*, 7(3), 2255. <https://doi.org/10.29140/ajal.v7n3.2255>
- Yamaoka, K. (2024). ChatGPT's motivational effects on Japanese university EFL learners: A qualitative analysis. *International Journal of TESOL Studies*, 6(3), 1–12. <https://doi.org/10.58304/ijts.20240303>

- Yıldız, C. (2024). ChatGPT integration in EFL education: A path to improved speaking self-efficacy. *Novitas-ROYAL (Research on Youth and Language)*, 18(2), 167–182. <https://doi.org/10.5281/zenodo.13861137>
- Zhang, X., Austin, B. W., & Ardasheva, Y. (2023). New directions in L2 self-efficacy research: Comparing bifactor and second-order models in the English public speaking domain. *Journal of Psycholinguistic Research*, 52(2), 569–587. <https://doi.org/10.1007/s10936-022-09910-7>