www.ijte.net

Beyond Theory: A Mixed-Methods Exploration of Postgraduate Engagement with ChatGPT for English Speaking Proficiency

Huanhuan Zhang
Macao Polytechnic University, China

Yujie Su
Wenzhou University, China

Xiaosu Xu
Wenzhou University, China

Vivian Ngan-Lin Lei
Macao Polytechnic University, China

Shanshan Hao
Macao Polytechnic University, China

To cite this article:

Zhang, H., Su, Y., Xu, X., Lei, V.N.L., & Hao, S. (2025). Beyond theory: A mixed-methods exploration of postgraduate engagement with ChatGPT for English speaking proficiency. *International Journal of Technology in Education (IJTE), 8*(4), 1022-1047. https://doi.org/10.46328/ijte.1221

International Journal of Technology in Education
2025, Vol. 8, No. 4, 1022-1047 https://doi.org/10.46328/ijte.1221

Beyond Theory: A Mixed-Methods Exploration of Postgraduate Engagement with ChatGPT for English Speaking Proficiency

Huanhuan Zhang, Yujie Su, Xiaosu Xu, Vivian Ngan-Lin Lei, Shanshan Hao

Article Info

Article History

Received:
5 December 2024
Accepted:
28 August 2025

Keywords

English speaking proficiency
ChatGPT
Higher education
Learning interest
EFL

Abstract

This study investigates the growing interest in ChatGPT's role in education, with a particular focus on its effects on postgraduate students' English-speaking proficiency—an area predominantly explored through theoretical perspectives with limited empirical evidence. The research aims to address this gap by examining the experiences and perceptions of Chinese postgraduate students using ChatGPT for IELTS speaking practice. Adopting a mixed-methods approach, the study integrates experimental engagement with ChatGPT, analysis of reflective diaries, and responses from questionnaires to evaluate the platform's efficacy in enhancing enthusiasm for English language practice. The findings reveal a positive impact of ChatGPT on learners' engagement, demonstrated by an increase in practice frequency. A significant correlation was observed between regular interaction with ChatGPT and improvements in English grammar skills among participants. Thematic analysis of reflective diaries underscored ChatGPT's contribution to language learning and students' proactive approach to acquiring English-speaking skills. However, there was a consensus among students regarding the need for more human-like interactions from ChatGPT. This study provides valuable insights into the potential of chatbots to enhance English-speaking skills in EFL contexts, offering practical recommendations for educators and developers to optimize chatbot learning experiences.

Introduction

Generative AI refers to a class of artificial intelligence models proficient in creating new data by identifying patterns within existing datasets. These models can generate content across various domains, including text, images, and music (Anantrasirichai & Bull, 2022; Naumova, 2023; Rudolph, Tan, & Tan, 2023). By leveraging deep learning techniques and neural architectures, these models can emulate human-like content creation. An example is ChatGPT, launched in 2020 by OpenAI, which performs multiple natural language processing tasks such as comprehension, text generation, and translation. Specifically designed for generating conversational responses, ChatGPT has undergone extensive training to facilitate human-like interactions (Goar, Yadav, & Yadav, 2023) . The versatility of ChatGPT enables its application across diverse fields, enhancing its distinctiveness (King, 2023; Thorp, 2023).

In the field of higher education, the rise of AI chatbots, particularly ChatGPT, can be attributed to their interactive capabilities, easy accessibility, and prompt feedback mechanisms (Abbas, Whitfield, Atwell, Bowman, Pickard, & Walker, 2022). These chatbots seamlessly integrate with conventional teaching practices, bolstering student engagement and fostering an environment conducive to self-directed learning tailored to the diverse needs of students (Ali et al., 2025; Haristiani & Rifai, 2021; Hamzah & Sani, 2021). ChatGPT, in particular, stands out for its ability to deliver personalized learning experiences by analyzing individual student data, thus elevating both their engagement and academic achievements (Hamzah, Li, & Sani, 2023). Educators can utilize ChatGPT's capabilities to gain insights into lesson planning, effective teaching methodologies, and classroom management (Kasneci et al., 2023). Additionally, as an online tutoring tool, ChatGPT offers customized support based on each student's distinct requirements, ensuring enhanced academic performance and deeper engagement (Pardos & Bhandari, 2023).

Research on AI chatbots in language learning has shown their positive impact, with unique technological benefits such as timeliness, personalization, and ease of use enhancing student communication in target languages (Mahmoud, 2022; Wahyuni, 2022). These chatbots offer tailored language resources based on prior interactions, improving student engagement. Their integration across various platforms, from webpages to mobile apps, ensures seamless interaction, capitalizing on the primary advantages of online educational tools (Bower, 2017).

However, specific research on ChatGPT's application in English as a Foreign Language (EFL) remains limited. For instance, while ChatGPT has shown potential as a valuable assistant by offering tailored feedback on aspects like grammar, vocabulary, and pronunciation (Rudolph et al., 2023), comprehensive research into its actual benefits remains nascent. Publications on ChatGPT's influence in education surged in 2023, indicating its impact starting from late 2022 (Lo, 2023). While existing studies highlight AI chatbots' educational advantages, there is a need to explore learners' interactions with ChatGPT more deeply (Jeon, 2021).

Current research often falls short in several areas: a limited focus on speaking practice, neglect of postgraduate education settings, and inadequate assessment of learner engagement. These insufficiencies create a practical need for the current study, which aims to fill these gaps by providing a detailed critique of the existing literature and setting a stronger stage for the research questions and objectives. Addressing this research gap is crucial for understanding how ChatGPT can be used to support English-speaking practice and improve educational outcomes. Our study adopts a mixed-methods approach, integrating experiments, reflective diaries, and questionnaires to comprehensively examine Chinese postgraduate students' experiences with ChatGPT for English speaking. Furthermore, we assess its potential to heighten their enthusiasm towards IELTS speaking practice. We present two central research questions:

Q1: What are the postgraduates’ perceptions of using ChatGPT for practicing IELTS speaking?

Q2: Can ChatGPT improve postgraduates’ interest in practicing IELTS speaking?

Understanding the impact of technology, especially AI, on learning, requires considering both its functional and emotional facets. AI's role in mimicking real-life interactions in education emphasizes the Social Presence Theory, which advocates for the importance of a "human" touch in digital learning for increased engagement. This is

particularly relevant in language learning, which involves human-like nuances beyond just grammar. Additionally, new educational tools can cause a "novelty effect"—a surge in initial engagement. Determining the longevity of this engagement is vital to assess the tool's sustained benefits.

Addressing the research gap has significant pedagogical implications. By evaluating how ChatGPT can enhance speaking practice in EFL settings, this study aims to contribute to broader educational goals, such as improving language acquisition methodologies. Personalized learning approaches can be refined through insights gained from ChatGPT's tailored feedback, leading to more effective and engaging language instruction. Furthermore, the findings could inform policy decisions related to AI integration in education, ensuring that technological advancements are harnessed to benefit both students and educators.

Taking into account these educational theories and technological phenomena, it becomes crucial to dive deeper and understand how specific tools, such as ChatGPT, fit into this landscape and affect EFL education. This paper evaluates the literature concerning ChatGPT's role in EFL education, traditional IELTS speaking practices, the Social Presence Theory, and the novelty effect of emerging technology. Employing mixed methods, including reflective diaries and questionnaires, the study examines postgraduate students' perceptions of using ChatGPT for IELTS speaking practice. The findings aim to provide pivotal insights to guide future research and practical implications at the intersection of AI and EFL teaching.

.

Literature Review

The Application of ChatGPT in EFL Learning

The capacity of tech-driven learning resources to aid language acquisition has garnered notable interest from educators and experts across multiple disciplines, especially following the introduction of ChatGPT by OpenAI in November 2022. ChatGPT has been recognized for enhancing student achievement and nurturing critical thinking. It offers rapid and precise responses, granting students easy access to current and pertinent data, aiding their comprehension of various subjects(Javaid, Haleem, Singh, Khan, & Khan, 2023). Moreover, it tailors learning to each student's pace, boosting motivation while allowing educators to dedicate more time to interactive tasks and personalized feedback (Limo, 2023).

In language learning contexts, several studies have explored ChatGPT's impact. Han et al. (2023) introduced ChEDDAR, a dataset from EFL students' interactions with ChatGPT during essay revisions, showing that students often perceived AI as a peer. Xiao and Zhi (2023) found that students from a top Chinese international university valued ChatGPT for its instant feedback and its role in boosting autonomy and critical thinking. Baskara (2023) underscored ChatGPT's potential in EFL writing instruction, emphasizing its motivational human-like text generation while also noting design and ethical issues. Mohamed (2023) examined EFL faculty members' views at Northern Border University, revealing mixed opinions: some appreciated its accuracy, while others voiced concerns about misinformation and hindrance to critical thinking. Overall, it was seen as a promising supplement to EFL education.

However, while these studies highlight various benefits of ChatGPT, they often do so in isolation, without deeply integrating or synthesizing findings. For instance, while Han et al. (2023) focus on essay revisions, and Xiao and Zhi (2023) on autonomy and critical thinking, there is limited discussion on how these findings interrelate or how discrepancies might arise. This fragmented approach underlines the need for a more cohesive analysis that compares and contrasts different findings to highlight overarching trends and gaps in the research.

The application of ChatGPT in EFL learning offers numerous benefits, such as tailored content and activities, promoting individualized learning, and boosting student motivation (Sánchez, 2023). Wahyuni (2022) observed that ChatGPT created a learner-friendly environment, enhancing language practice. Furthermore, Kim and Kim (2021) reported significant improvement in EFL students' communication skills after practicing with ChatGPT. Kasneci et al. (2023) noted that ChatGPT aids learners in improving writing and vocabulary, offering tailored resources, and assisting in tasks like research and domain-specific language skills.

Despite these positive reports, critical evaluations of methodologies and potential biases are often missing. For example, while García Sánchez (2023) and Wahyuni (2022) report enhancements in motivation and language practice, the specific contexts and conditions under which these improvements occur are not thoroughly examined. Additionally, potential biases in self-reported data or the novelty effect influencing initial student enthusiasm are seldom addressed, which could skew the perceived effectiveness of ChatGPT.

Furthermore, there is a notable disconnect between the literature review and the specific research questions posed in this study. While various aspects of AI in language learning are mentioned, there is little direct connection to how these findings relate to postgraduate students' IELTS speaking practice. This gap highlights the need for a more targeted review that explicitly links existing research to the study's objectives.

Addressing the limitations and concerns mentioned in previous studies is crucial for the current research. For instance, Kohnke (2023) highlighted issues such as promoting cheating and assessment difficulties. García-Peñalvo (2023) and Ulla et al. (2023) pointed out ethical and pedagogical challenges, including the inconsistency and potential for plagiarism in ChatGPT's responses. Rudolph et al. (2023) mentioned "hallucinations" where ChatGPT delivers plausible yet incorrect responses, which can mislead students.

Given these challenges, there remains a scarcity of research delving into student views and experiences with ChatGPT, particularly in postgraduate settings. This gap is crucial for understanding the broader implications of ChatGPT in EFL education and for optimizing its use to benefit learners. By addressing this research gap, the current study aims to contribute to broader educational goals, such as improving language acquisition methodologies, refining personalized learning approaches, and informing policy decisions related to AI integration in education.

IELTS Speaking Training

The IELTS Speaking test assesses candidates based on fluency, vocabulary, grammar, and pronunciation.

Speaking proficiency is often deemed the most daunting language skill (Kürüm, 2016), traditionally addressed through standard methods, leaving many learners struggling to achieve true English fluency (West, 2015). The direct, face-to-face format of the IELTS Speaking test further intensifies these challenges (Dashti & Razmjoo, 2020).

Innovative approaches are being explored to enhance speaking skills. Han and Niu (2021) leveraged virtual scenarios, immersing learners in AI-driven, real-life simulations that foster spontaneous communication. Contrastingly, Santhanasamy and Yunus (2022) employed a flipped classroom, shifting the lecture paradigm. Students independently delved into materials, reserving classroom sessions for enriching, interactive speaking tasks, thereby amplifying engagement and self-driven motivation. Embracing technology further, Sherine (2020) harnessed WhatsApp for voice note exchanges and discussions, enhancing fluency and learner confidence. Similarly, Iio et al. (2019) deployed a language-learning robot, offering immediate feedback on pronunciation and refining learners' articulation skills.

Despite advances in innovative teaching strategies for IELTS Speaking, there remains a conspicuous gap in harnessing the potential of Natural Language Processing systems like ChatGPT. While traditional methods and newer technological approaches show promise, the integration of advanced AI tools like ChatGPT in IELTS Speaking preparation is underexplored. This study aims to fill this gap by investigating how ChatGPT can augment and elevate the IELTS Speaking preparation experience, particularly for postgraduate students.

Social Presence Theory

Three categories of social presence were identified in chatbot-supported language learning: interpersonal communication (e.g., students' self-discourse), open communication (e.g., continuing a thread, asking questions, expressing agreement), and cohesion communication (e.g., using vocatives and greetings). Xu and Warschauer (2020) suggested that interpersonal communication can be established through students' self-disclosure with chatbots. Moreover, using a chatbot as a learning partner can enhance social interaction by exchanging self-disclosure information. Zaghlool et al. (2023) showed that chatbots could continue a thread with students by providing examples of possible responses when students faced difficulties answering a question (Lin & Chang, 2020). Embracing chatbots in language learning can encourage an open learning climate of interpersonal communication, helping students overcome nervousness about speaking the target language and promoting their willingness to communicate (Ayedoun et al., 2015). This approach helps students better understand learning objectives, supports collaborative learning (Tegos, Demetriadis, & Tsiatsos, 2014), and strengthens their sense of social presence within virtual language environments (Wang, Petrina, & Feng, 2017).

However, the immature development of chatbot technology can lessen social presence in students' language learning. Fryer, Ainley, Thompson, Gibson, & Sherlock (2017) criticized chatbots for their inability to maintain students' language learning interest; students' interest in speaking tasks with chatbot partners dropped after the first task compared with speaking to a human partner. Similarly, Hsu (2020) found that participants who interacted face-to-face with human interlocutors perceived higher socialization than those who conversed with chatbot

partners in a second language context. These findings suggest that while chatbots can initially boost engagement, their long-term effectiveness in maintaining social presence is questionable.

The Novelty Effect of New Technology

The "novelty effect" describes how new technology can temporarily enhance student motivation and performance due to its novelty. Chen et al. (2016) found that the introduction of new technology spurred student engagement, leading to improved performance, not necessarily because of the technology's intrinsic value, but because it was exciting and new. In a 16-week study, Fryer et al. (2017) noticed this effect with chatbots; initial student interest waned after the first task as the novelty of the chatbot diminished. Similarly, Gallacher et al. (2018) noted that students viewed chatbots as novelties rather than consistent tools for language practice. Addressing these phenomena is crucial for understanding the broader implications of using AI tools like ChatGPT in language education. By examining how ChatGPT influences student engagement and learning outcomes, particularly in the context of IELTS speaking practice, this study aims to provide comprehensive insights that can optimize the use of AI in EFL education.

In sum, the review addresses the potential and challenges of using ChatGPT in EFL learning, explores innovative methods in IELTS speaking training, emphasizes the significance of social presence theory, and examines the impact of the novelty effect. Further research in these areas can offer comprehensive insights into optimizing AI tools like ChatGPT for language education. By synthesizing existing findings and critically evaluating methodologies, this study aims to bridge the gap between current research and practical applications, ultimately contributing to broader educational goals and theoretical advancements in EFL education.

Method

Participants and Context

This study involved 48 first-year postgraduate students from a comprehensive university in China, with diverse majors such as Engineering and Psychology, none of which were English-related. The group consisted of 17 males and 31 females. Participants were from varying cognitive and linguistic backgrounds, which are essential for understanding their starting points in the study. As part of a nine-week IELTS preparation course, students had the optional opportunity to use ChatGPT 3.5 for extra speaking practice. While participation was unpaid, it provided added IELTS prep support. The effectiveness of ChatGPT 3.5 was gauged via questionnaires and reflective diaries assessing speaking proficiency. Necessary consents and approvals, including one from the university's IRB (SFS2023), were obtained, and the study strictly followed ethical guidelines.

Research Design

The study employed a mixed-methods design to thoroughly investigate the impact of ChatGPT on students' speaking proficiency and engagement. This methodological approach was selected to align with the study's objectives, which include assessing both quantitative improvements in speaking skills and gaining qualitative

insights into students' experiences. By integrating questionnaires and reflective diaries, the research facilitates a comprehensive exploration of the research questions, offering both numerical data and personal reflections. This multifaceted approach enhances the robustness of the findings, capturing the nuanced effects of ChatGPT on language learning.

The experimental setup included specific details about the settings in which ChatGPT was used and the interaction modes. ChatGPT was employed in both textual and vocal interaction modes to provide comprehensive practice opportunities. Its role in speaking practice involved providing feedback on pronunciation, vocabulary usage, and fluency during simulated IELTS speaking tasks. The settings encompassed both classroom environments and individual practice sessions, ensuring diverse interaction contexts. This detailed approach allowed for a thorough examination of the varied ways in which ChatGPT influenced students' language learning experiences.

Instruments

Questionnaire

The survey instrument used in this study was inspired by the works of Hong, Hwang, Liu, Ho, & Chen (2014), Tsai (2015), and Wang et al. (2023) to gather data on the effects of ChatGPT on learners' English-speaking proficiency and learning interests. The questionnaire was divided into four sections. The first section collected demographic information, including gender, grade, and major. The second section examined the frequency and duration of students' usage of ChatGPT for oral training during the IELTS course. The third section assessed the impact of ChatGPT on students' speaking proficiency, covering aspects such as vocabulary, sentence patterns, grammar, language organization, and fluency. The fourth section explored the influence of ChatGPT on students' interest in spoken English learning, focusing on aspects like liking, enjoyment, and engagement.

The questionnaire contained a total of 19 items, with the third and fourth sections utilizing a Likert scale with five response options ranging from "Strongly Disagree" to "Strongly Agree. The questionnaire was administered using the Wenjuanxing platform, which required participants to answer all questions before submission. This mandatory response format ensured that there were no missing data in the dataset. By enforcing complete responses, the analysis could proceed without concerns about missing data, providing a comprehensive and accurate evaluation of the participants' experiences and outcomes. The development and validation of the questionnaire followed established protocols to ensure its reliability and validity, providing a robust tool for measuring the study's key variables.

Reflective Diary

Participants were instructed to write two reflection diaries, one in week three and another in week seven, to capture their evolving experiences and insights as they became more familiar with using ChatGPT for their oral training. These reflection diaries consisted of six open-ended questions, allowing participants to express their thoughts, challenges, and progress regarding their utilization of ChatGPT.

Data Analysis

The collected data from questionnaires and diaries will be analyzed using both quantitative and qualitative methods to address the research questions. Quantitative data from the questionnaires will be statistically analyzed to identify trends and measure improvements in speaking proficiency. Qualitative data from the reflective diaries will be thematically analyzed to extract insights into students' experiences, challenges, and perceived benefits of using ChatGPT. This dual approach ensures a comprehensive understanding of the impact of ChatGPT on IELTS speaking practice.

Results

Results of the Questionnaire

The SPSS-based reliability analysis showed Cronbach's alpha values of 0.93 for the English ability questionnaire and 0.96 for the learning interests questionnaire. Both values exceed the threshold of 0.7, indicating strong internal consistency among the questions. A Spearman correlation analysis was used to investigate the link between the frequency and duration of ChatGPT use and factors like English proficiency and interest in learning.

Table 1 and 2 showcase the notable enhancements in students' English proficiency attributed to their practice with ChatGPT.

Table 1. Enhancement of Students' English Proficiency Relative to Varied Duration of ChatGPT Usage

Questions		0-10 days (M±SD)	10 - 20days (M±SD)	20 - 30days (M±SD)	More than 30 days (M±SD)	Total (M±SD)
Q1	ChatGPT helps improve my English vocabulary.	4.12±0.81	3.93±0.80	4.14±0.90	4.40±0.70	4.13±0.79
Q2	ChatGPT helps to enrich my English sentence patterns.	4.06±0.93	4.13±0.83	4.00±0.82	4.40±0.70	4.15±0.83
Q3	ChatGPT helps to improve my English grammar and thus improve the accuracy of my speech.	4.00±0.89	4.40±0.63	4.14±0.90	4.4±0.70	4.23±0.78
Q4	ChatGPT helps improve my logic and language organization skills.	4.13±0.96	3.93±0.80	4.29±0.76	4.40±0.70	4.15±0.83

Questions		0-10 days (M±SD)	10 - 20days (M±SD)	20 - 30days (M±SD)	More than 30 days (M±SD)	Total (M±SD)
Q5	The multiple rounds of conversation feedback and assessment provided byChatGPT helped improve my spoken English skills and fluency.	4.12±0.81	3.80±0.86	4.14±0.69	4.50±0.71	4.10±0.81
Total		4.09±0.83	4.04±0.61	4.14±0.77	4.42±0.65	4.15±0.71
N		16	15	7	10	48

Table 2. Progress in Students' English Proficiency Relative to Different Daily Usage Durations of ChatGPT

Questions		0-10min (M±SD)	10-20min (M±SD)	20-30min (M±SD)	More than 30min (M±SD)	Total (M±SD)
Q1	ChatGPT helps improve my English vocabulary.	4.00±0.73	4.11±0.88	4.13±0.84	4.60±0.55	4.13±0.79
Q2	ChatGPT helps to enrich my English sentence patterns.	3.94±0.68	4.16±0.9	4.25±1.04	4.60±0.55	4.15±0.83
Q3	ChatGPT helps to improve my English grammar and thus improve the accuracy of my speech.	3.94±0.68	4.26±0.87	4.50±0.76	4.60±0.55	4.23±0.78
Q4	ChatGPT helps improve my logic and language organization skills.	4.13±0.72	3.89±0.94	4.50±0.76	4.60±0.55	4.15±0.83
Q5	The multiple rounds of conversation feedback and assessment provided byChatGPT helped improve my spoken English skills and fluency.	4.06±0.68	4.00±0.82	4.13±1.13	4.60±0.55	4.10±0.81
Total		4.01±0.61	4.08±0.79	4.30±0.76	4.60±0.55	4.15±0.71
N		16	19	8	5	48

Table 1 and 2 depict how students who practiced with ChatGPT noted improvements in areas such as English

vocabulary, sentence construction, grammar, language structure, and speaking skills, with scores exceeding 4 (agree). This indicates a widespread acknowledgment of ChatGPT's effectiveness as a learning tool. Among all the facets, Q3, which states "ChatGPT aids in enhancing my English grammar, leading to more accurate speech," garnered the highest rating with an average score of 4.23.

Table 1 further showcases a rising trend in students recognizing improvements in their English proficiency with increased practice days using ChatGPT. This trend moves from a starting score of 4.04 (for 10-20 days) to a more elevated score of 4.42 (for more than 30 days). Similarly, Table 2 demonstrates that as students dedicate more daily time to practicing with ChatGPT, they also report heightened recognition of their English advancements. The appreciation grows from a score of 4.01 (for 0-10 minutes daily) to a heightened rating of 4.60 (for over 30 minutes daily), underscoring the beneficial correlation between practice duration and English skill enhancement.

Table 3. Spearman Correlation Analysis of ChatGPT Practice Frequency and Students' English Ability

Questions		Length of practice		Daily practice time	
		ρ	Sig. (Double tail)	ρ	Sig. (Double tail)
Q1	ChatGPT helps improve my English vocabulary.	0.11	0.457	0.172	0.241
Q2	ChatGPT helps to enrich my English sentence patterns.	0.093	0.528	0.279	0.055
Q3	ChatGPT helps to improve my English grammar and thus improve the accuracy of my speech.	0.154	0.297	.335*	0.02
Q4	ChatGPT helps improve my logic and language organization skills.	0.101	0.494	0.193	0.189
Q5	The multiple rounds of conversation feedback and assessment provided by ChatGPT helped improve my spoken English skills and fluency.	0.151	0.304	0.164	0.265
Total		0.138	0.35	0.232	0.113
N		48		48	

** At level 0.05 (two-tailed), the correlation was significant.*

Table 3 reveals that there is a positive correlation between Q3 (which focuses on the improvement of English grammar and the resulting enhancement in speech accuracy using ChatGPT) and the amount of time students spend practicing with ChatGPT daily.

However, the correlation coefficients are relatively low, suggesting a weak relationship. Regarding learning interests, Table 4 provides a detailed view of how different durations of ChatGPT usage affect students'

enthusiasm for learning. This table breaks down the enhancement in students' learning interests into three categories - 'Like', 'Enjoyment', and 'Engagement' - over various periods of ChatGPT use.

Table 4. Enhancement in Students' Learning Interests Correlating with Various Durations of ChatGPT Usage

Questions		0-10 days (M±SD)	10-20days (M±SD)	20-30days (M±SD)	More than 30 days (M±SD)	Total (M±SD)
Q1	After my last use of the ChatGPT, I still liked using it.	4.12±0.81	4.2±0.56	4.14±0.9	4.4±0.84	4.21±0.74
Q2	After my last use of the ChatGPT, I am still eager to use it again.	4.06±0.85	4.07±0.59	4.14±0.9	4.5±0.71	4.17±0.75
Q3	Even though I didn't score well the last time I used the ChatGPT, I still want to use it.	4.06±0.85	4±0.85	4±0.82	4.4±0.84	4.1±0.83
LIKE		4.08±0.83	4.09±0.6	4.1±0.81	4.43±0.77	4.16±0.74
Q4	The last time I used the ChatGPT for IELTS speaking, I still felt excited.	3.81±0.91	4.07±0.7	4.14±0.9	4.2±0.92	4.02±0.84
Q5	The last time I used the ChatGPT for IELTS speaking, I still felt focused.	3.88±0.81	4.2±0.56	4.14±0.9	4.5±0.71	4.15±0.74
Q6	The last time I used the ChatGPT for IELTS speaking, I still felt enjoyed.	3.87±0.72	4.13±0.52	4.14±0.9	4.3±0.82	4.08±0.71
ENJOYMENT		3.85±0.78	4.13±0.56	4.14±0.9	4.33±0.77	4.08±0.73
Q7	The last time I used the ChatGPT for IELTS speaking, I lost track of time.	3.75±0.78	3.93±0.8	4.14±0.9	4.4±0.84	4±0.83
Q8	The last time I used the ChatGPT for IELTS speaking, I still felt focused.	3.75±0.78	4.07±0.59	4.29±0.76	4.3±0.82	4.04±0.74

Questions		0-10 days (M±SD)	10-20days (M±SD)	20-30days (M±SD)	More than 30 days (M±SD)	Total (M±SD)
Q9	The last time I used the ChatGPT for IELTS speaking, I still felt attracted.	3.81±0.75	3.93±0.7	4.14±0.9	4.3±0.82	4±0.77
ENGAGEMENT		3.77±0.74	3.98±0.61	4.19±0.84	4.33±0.79	4.01±0.73
Total		3.92±0.72	4.05±0.56	4.13±0.86	4.36±0.71	4.08±0.69
N		16	15	7	10	48

Table 4 demonstrates that as students spent more days practicing with ChatGPT, there was a notable uptick in their scores on the learning interest scale. This is evident in the 'Like', 'Enjoyment', and 'Engagement' dimensions. In the 'Like' dimension, average scores surpassed 4 (agree) across varied ChatGPT usage durations, pointing to a positive inclination towards practicing with ChatGPT. For the 'Enjoyment' dimension, students practicing between '0-10' days reported an average score of 3.85, but this mean score rose above 4 (agree) once they practiced beyond 10 days. Similarly, for the 'Engagement' dimension, while the score averaged at 3.77 for '0-10 days', it increased to 3.98 for '10-20 days' and surpassed the 4 (agree) mark for practice durations exceeding 20 days.

Table 5 sheds light on how varying durations of daily ChatGPT usage correlate with students' learning interests. These interests are evaluated based on three criteria: 'Like', 'Enjoyment', and 'Engagement'.

Table 5. Enhancement of Students' Learning Interests in Relation to Varied Daily Usage Durations of ChatGPT

Questions		0-10min (M±SD)	10 - 20min (M±SD)	20 - 30min (M±SD)	More than 30min (M±SD)	Total (M±SD)
Q1	After my last use of the ChatGPT, I still liked using it.	4.19±0.75	4.21±0.71	4.13±0.84	4.4±0.89	4.21±0.74
Q2	After my last use of the ChatGPT, I am still eager to use it again.	4.13±0.81	4.11±0.74	4.25±0.71	4.4±0.89	4.17±0.75
Q3	Even though I didn't score well the last time I used the ChatGPT, I still want to use it.	4±0.82	4.05±0.91	4.13±0.84	4.6±0.55	4.1±0.83
LIKE		4.1±0.77	4.12±0.74	4.17±0.76	4.47±0.77	4.16±0.74
Q4	The last time I used the ChatGPT for	3.94±0.93	4±0.75	4±0.93	4.4±0.89	4.02±0.84

Questions		0-10min (M±SD)	10 - 20min (M±SD)	20 - 30min (M±SD)	More than 30min (M±SD)	Total (M±SD)
	IELTS speaking, I still felt excited.					
Q5	The last time I used the ChatGPT for IELTS speaking, I still felt focused.	4±0.73	4.11±0.74	4.38±0.74	4.4±0.89	4.15±0.74
Q6	The last time I used the ChatGPT for IELTS speaking, I still felt enjoyed.	3.88±0.72	4.05±0.62	4.38±0.74	4.4±0.89	4.08±0.71
ENJOYMENT		3.94±0.76	4.05±0.67	4.25±0.77	4.4±0.89	4.08±0.73
Q7	The last time I used the ChatGPT for IELTS speaking, I lost track of time.	4±0.73	3.84±0.83	4.13±0.99	4.4±0.89	4±0.83
Q8	The last time I used the ChatGPT for IELTS speaking, I still felt focused.	3.81±0.75	4±0.67	4.25±0.89	4.6±0.55	4.04±0.74
Q9	The last time I used the ChatGPT for IELTS speaking, I still felt attracted.	3.94±0.68	3.89±0.74	4.13±0.99	4.4±0.89	4±0.77
ENGAGEMENT		3.92±0.68	3.91±0.67	4.17±0.94	4.47±0.77	4.01±0.73
Total		3.99±0.68	4.02±0.64	4.18±0.8	4.42±0.85	4.08±0.69
N		16	19	8	5	48

Table 5 demonstrates a clear positive trend, showing that as students increase their daily ChatGPT usage time, their scores on the learning interest scale also rise. This pattern is consistent across the 'Like', 'Enjoyment', and 'Engagement' dimensions. In the 'Like' category, the average score consistently surpasses 4 (agree) for all durations of daily ChatGPT use, underscoring the positive influence of consistent ChatGPT practice on students' liking levels. For 'Enjoyment', the average score is 3.94 for those practicing 0-10 minutes daily, with scores exceeding 4 (agree) when practice extends beyond 10 minutes each day. Regarding 'Engagement', the score averages 3.92 for 0-10 minutes of daily use, dips slightly to 3.91 for 10-20 minutes, but surpasses 4 (agree) for durations over 20 minutes daily.

Table 6 presents the findings from a Spearman correlation analysis that probes the link between the frequency of ChatGPT practice and students' learning interests across different dimensions.

Table 6. Spearman Correlation Analysis between ChatGPT Practice Frequency and Students' Learning Interests

Questions		Length of practice		Daily practice time	
		ρ	Sig. (Double tail)	ρ	Sig. (Double tail)
Q1	After my last use of the ChatGPT, I still liked using it.	0.12	0.41	0.04	0.77
Q2	After my last use of the ChatGPT, I am still eager to use it again.	0.19	0.21	0.09	0.53
Q3	Even though I didn't score well the last time I used the ChatGPT, I still want to use it.	0.12	0.42	0.16	0.27
LIKE		0.12	0.42	0.10	0.50
Q4	The last time I used the ChatGPT for IELTS speaking, I still felt excited.	0.17	0.24	0.10	0.49
Q5	The last time I used the ChatGPT for IELTS speaking, I still felt focused.	.291*	0.05	0.20	0.16
Q6	The last time I used the ChatGPT for IELTS speaking, I still felt enjoyed.	0.23	0.12	0.28	0.06
ENJOYMENT		0.21	0.15	0.18	0.22
Q7	The last time I used the ChatGPT for IELTS speaking, I lost track of time.	.301*	0.04	0.11	0.44
Q8	The last time I used the ChatGPT for IELTS speaking , I still felt focused.	.302*	0.04	.309*	0.03
Q9	The last time I used the ChatGPT for IELTS speaking, I still felt attracted.	0.24	0.10	0.15	0.32
ENGAGEMENT		.295*	0.04	0.19	0.21
Total		0.21	0.15	0.17	0.26
N		48		48	

** At level 0.05 (two-tailed), the correlation was significant.*

Table 6 unveils results from the Spearman correlation analysis, showing positive associations between Q5 (The last time I used ChatGPT, I remained focused), Q7 (The last time I used ChatGPT, I lost track of time), and Q8 (The last time I used ChatGPT, I remained focused) with the 'Engagement' dimension related to the number of

ChatGPT practice days. Furthermore, Q8 also demonstrates a positive linkage with the duration of daily ChatGPT practice.

In sum, the reliability analysis confirmed robust internal consistency for the English ability and learning interests questionnaires. Students using ChatGPT regularly observed noticeable improvements in English proficiency, particularly in grammar, which translated to more accurate speech. A positive correlation emerged between the duration of daily ChatGPT practice, perceived advancements in English, and heightened learning interests across several dimensions. Increased practice with ChatGPT consistently aligned with elevated learning enthusiasm and engagement among students.

Thematic Analysis of the Reflective Diary

Data transcribed from students' reflective journals were saved sequentially in the "Students' Reflective Journal" folder using Microsoft Word. Out of all, 48 students provided their journals, each addressing six specific questions. Following the coding steps proposed by Braun and Clarke (2006), the first author reviewed the content, creating initial codes and grouping some into potential themes. Both researchers then independently conducted a thematic analysis. After collaborative discussions about the codes and themes, they reached a consensus on the final set. Due to space constraints, the following discussion will briefly outline the identified themes and codes Braun and Clarke (2006).

Q1: Do you like using ChatGPT as your private IELTS speaking tutor? Why or why not?

Out of 48 participants, 38 favored interacting with ChatGPT, highlighting its efficiency, proficiency, and convenience. They believed that the chatbot alleviated feelings of nervousness and created a non-judgmental, relaxed environment. In contrast, two participants felt that ChatGPT couldn't emulate human-like empathy and social interaction. This feedback mirrors Ayedoun, Hayashi, and Seta.'s (2015) research, which underscored chatbots' potential to diminish anxiety and bolster social engagement. However, Fryer et al. (2017, 2020), Fryer, Nakao, and Thompson (2019), and Hsu (2020) pointed out the inherent limitations of chatbots, including a diminished social presence and lack of maturity, which influenced some users' preference for human interaction. Additionally, the necessity to type when communicating with the chatbot further diminished the sense of genuine social interaction.

Q2: Do you prefer practicing IELTS speaking in the classroom with a human or with ChatGPT? Why?

Out of 48 participants, 14 strongly favored practicing speaking with ChatGPT, whereas 18 leaned towards real-person interactions in class. Five students experienced a change in preference, with four transitioning from ChatGPT to human interactions and one the opposite; one student had ambivalent feelings about ChatGPT. Those inclined towards real-person practice emphasized the value of genuine linguistic and non-linguistic interactions, noting enhanced focus, efficacy, and comprehensive feedback. Nevertheless, they recognized ChatGPT's merits in offering a relaxed practice setting and assistance in emulating conversational exercises.

Q3: How has ChatGPT helped you with your IELTS speaking skills? (Such as expanding vocabulary? Broadening thinking? ...)

Upon examining the reflective diaries, we discerned six principal themes associated with the benefits of employing ChatGPT for IELTS speaking practice, which further branched into sixteen codes:

Expanding Vocabulary

For many English learners, acquiring vocabulary presents a significant challenge. This encompasses not only the meanings of words but also the related contextual information (Harris, Golinkoff, & Hirsh-Pasek, 2011; Hwang, Mancilla-Martinez, McClain, Oh, & Flores, 2020). Additionally, a robust vocabulary foundation is crucial for fluent expression and the subsequent understanding and acquisition of grammar (Liu, 2021; Xu, 2022).

Participants generally offered favorable feedback on how ChatGPT assisted their English vocabulary acquisition, for instance,

> *"It refines my vocabulary, makes my expressions more sophisticated, provides multiple reasonable word choices in a different contest, and make my language use less monotonous." (Zhang, Q.J.)*

They also conveyed that utilizing ChatGPT enhanced their comprehension of English vocabulary, for instance,

> *"I surprisingly discover that some words are actually more native and accurate than other words. I find that my ability to read English papers is also improving." (Fang, Y.L.)*

Broadening Thinking

Critical thinking encompasses aspects such as open-mindedness (Facione, 2000), and receptivity to others' viewpoints (Halpern, 1998). Participants provided feedback on varied thinking perspectives, adaptability in thought processes, depth of reasoning, and inventive thinking. For example,

> *"ChatGPT expands more dimensions to think about questions, get more knowledge from the conversation." (S36) "It answers my question comprehensively" (Huang, P.X.) "In addition to convenience, ChatGPT provides us with more ideas than we can imagine." (Zhou, X.F.)*

Expressiveness

Expressiveness is characterized by vibrant, impactful, and compelling communication, whether in speech or artistic presentations, encompassing attributes like articulacy, eloquence, and expressivity (Gray, 2014) (American Heritage® Dictionary of the English Language 2011). Participants noted an enhanced ability to articulate feelings and ideas when using ChatGPT as opposed to human interactions. For example,

> *"I can express my inner thoughts more vividly and make the communication more interesting." (Tong, D.) "When I practice my spoken English with ChatGPT, my vocabularies and sentences start to become more abundant and natural, in my view." (Fang, Y.L.) "I struggle with fluency and coherence, and ChatGPT has been great at pointing out these issues and providing suggestions on how to improve them." (Wang, J.L.)*

Enhancing Grammar

English grammar serves as a foundational component that directly impacts both expression and comprehension across all language facets, such as listening, speaking, reading, and writing. The findings of the research aligns with Pan and Zhou's (2018) research, which suggests that utilizing computer technology to learn grammar can bolster grammatical understanding. For instance,

> *"ChatGTP can help me improve my Grammar and sentence structure so that I can express my ideas more accurately." (Zhao, Z.Y.)*

Improving Pronunciation

Typically, Chinese students learn English in non-English settings, which often results in limited practical application of their knowledge. However, with computer-assisted technology, they can identify pronunciation errors and align with the accurate, objective, and standardized pronunciation guidance provided. This observation of this research is supported by studies conducted by Xiao and Park (2021), Han and Huwan (2022), and Xu (2022). For example,

> *"I can also assist with pronunciation and stress patterns by modeling correct pronunciation and providing practice exercises." (Wang, J.K.) "First of all is Pronunciation, ChatGTP can help me practice correct pronunciation, including vowels, consonants, stress, and intonation. I can improve my pronunciation by repeating sentences and words with ChatGTP." (Zhao, Z.Y.)*

Saving Time

In addition to the language-related insights provided by ChatGPT, as highlighted in the reflective diaries, students also emphasized its contribution to more efficient learning.

> *"If I send an inquiry of a word, it can rapidly locate and list a thesaurus of this word, which not only expands my vocabulary but also save my time in acquisition." (Xiao, W.J.)*

Q4: What are the main problems you face with your IELTS speaking? Can ChatGPT resolve these issues for you?

Reflective diary analysis identified five key challenges students faced in IELTS speaking, with pronunciation and grammar being paramount. Students mentioned ChatGPT's inability to correct pronunciation in real-time, as it lacks audio input, but believed a software update could remedy this. For grammar, they viewed ChatGPT as a valuable guide, pinpointing errors and suggesting fixes. Other challenges encompassed limited vocabulary, struggles with fluid speech, and restricted knowledge breadth. Despite these hurdles, students recognized ChatGPT as an advantageous learning asset. For instance,

> *"ChatGPT can promptly correct my grammatical mistakes, and at the same time offer well-accepted expressions." (Xiao, W.J.) "ChatGTP helped me practice using different sentence structures and correct grammar by repeating sentences and words with it." (Zhao, Z.Y.)*

Q5: By using ChatGPT, what changes have you noticed in your attitude, confidence, and motivation towards learning English speaking?

Upon analyzing the reflective journals, participants' changes were categorized into two primary themes, further divided into ten subthemes.

Fulfilling Interpersonal Communication

Interpersonal communication, a category of social presence, encompasses two essential concepts: intimacy and immediacy (Gunawardena, 1995; Short, Williams, & Christie, 1976). Participants remarked that ChatGPT fostered a personalized practice setting, solely involving the users. This created a relaxed and congenial interaction devoid of social anxieties.

Furthermore, ChatGPT's capability to respond instantly amplified these feelings of intimacy and immediacy. Such immediate and intimate interactions with ChatGPT encouraged students to participate more frequently and eagerly in its natural language dialogues, mirroring findings from Shoufan (2023). This engagement subsequently led to increased self-disclosure and heightened enthusiasm for English learning. For instance,

> *"Practicing speaking with ChatGPT helped relieve their anxiety and social phobia symptoms. ChatGPT provided wider choices of spoken topics and simulated interpersonal communication encountered in daily life." (XiaoWJ)*

Motivating Open Communication

Open communication, another facet of social presence, incorporates various subjective elements tied to conversational participants, including attitudes, confidence, and motivation. Notably, confidence has been identified to have multiple mediating roles between self-efficacy and English-speaking anxiety (Demirdöken & Okur, 2022). Participants revealed a boost in their confidence after utilizing ChatGPT over a duration, echoing findings from Sánchez-Ruiz, Moll-López, Nuñez-Pérez, Moraño-Fernández, & Vega-Fleitas (2023). This

suggests that consistent interaction with ChatGPT can positively influence learners' self-confidence in speaking English.

> *"I felt that through it (ChatGPT) I have increased my self-confidence and expressed my inner thoughts more fluently and smoothly." (Wang, H.H.)*

Q6: What difficulties have you encountered when using ChatGPT for IELTS speaking practice? How did you solve them?

Upon analysis of the reflective journals, researchers identified three predominant themes along with twelve subthemes that underscored the limitations or drawbacks of using ChatGPT for IELTS speaking preparation.

English Proficiency

Participants noted challenges in their interactions with ChatGPT due to their own limited English proficiency. For students with weaker English skills, effectively utilizing ChatGPT proved to be a challenge. Their uncertainties in framing correct expressions often led to struggles in obtaining accurate responses from the chatbot. For example,

> *"In some cases, it didn't accurately recognize my responses, resulting in a lack of accuracy in the ChatGPT's feedback on my responses." (ChenJT) "Sometimes the answer to a question that I cannot express clearly may not be what I expected, and this requires me to constantly adjust the details and focus of the question to guide it to provide the answer I need." (Zhu, M.C.)*

Pronunciation Guidance

Many participants expressed a lack of confidence in their English proficiency, particularly in pronunciation. As a result, the majority felt that their interactions with ChatGPT did not yield the anticipated responses, attributing this to potential misunderstandings stemming from their language skills. For instance,

> *"At the beginning of usage, ChatGPT actually did not have the feature of oral conversation." (Li, J.Y.) "It cannot provide pertinent feedback on my oral English." (Wang, K.F.)*

Limit of Use

The majority of issues reported pertained to technological limitations. The version of ChatGPT that participants used was tailored for Chinese users, which may have led to some of the challenges they encountered. For instance,

> *"But it is a little bit hard to use. For example, I have to frequently change my VPN location to avoid being banned by OpenAI. Additionally, ChatGPT only assists me with text-based interactions, and I have*

not discovered a way to utilize its pronunciation capabilities." (Ma, X.G.) "I also encountered some technical difficulties in the process, such as unstable website, which I overcame through other software unstable website" (Le, Y.Q.) "It was difficult to register on OpenAI's ChatGPT platform. And I do not know how to properly listen and speak with ChatGPT." (Yu, Z.N.)

In sum, the study on ChatGPT as an IELTS Speaking Tutor revealed mixed sentiments among student participants. Many students praised ChatGPT for its proficiency, efficiency, and ability to create a relaxed learning environment. However, there were concerns regarding its inability to emulate human empathy and genuine social interaction. While a significant number of students noted improvements in vocabulary and grammar after using ChatGPT, the platform's lack of audio functionality posed challenges, particularly in pronunciation guidance.

Discussion and Conclusion

This study utilized both questionnaire data and thematic analysis of student reflective diaries to explore postgraduates' perceptions of using ChatGPT for speaking practice in an IELTS setting. The questionnaire revealed several perceived patterns: students commonly felt that frequent interactions with ChatGPT might enhance their grasp of various English proficiency areas such as vocabulary, sentence structures, and grammar. However, it is essential to note that these sentiments are based on personal evaluations and may not reflect objective advancements.

The current findings align with previous research indicating that AI chatbots can positively impact language learning (Han et al., 2023; Xiao & Zhi, 2023). Prior studies such as Baskara (2023) and Mohamed (2023) highlighted motivational benefits and concerns about misinformation. Our study provides a nuanced view where students appreciated ChatGPT for improving grammar but found its help with pronunciation limited. These discrepancies might be due to differences in study designs, contexts, or participant backgrounds.

The thematic analysis of students' self-reports suggests that interactions with AI chatbots positively influenced their speaking performance and heightened their interest in learning English. Many students noted perceived enhancements in vocabulary and grammar, corroborating the quantitative results. However, qualitative insights also revealed that students experienced limited gains in pronunciation due to the lack of non-verbal feedback. This qualitative dimension supports the quantitative findings by offering a deeper understanding of the improvements and areas needing more attention. Direct quotes from the reflective diaries, such as "I feel more confident speaking English now" and "ChatGPT helps me with grammar, but I still struggle with pronunciation," highlight these themes and provide a richer understanding of participant experiences.

One limitation of this study is the relatively small sample size of 48 participants, which may limit the generalizability of the findings. Additionally, the study focused solely on the impact of ChatGPT on IELTS speaking proficiency, which cannot predict its effects on other language skills or aspects of language learning. These limitations suggest caution in interpreting the findings and underscore the need for further research with larger, more diverse samples.

While the study provides preliminary insights into the potential benefits of using ChatGPT for IELTS speaking practice, some explanations remain speculative. For instance, the correlation between increased use of ChatGPT and perceived improvements in grammar might be due to the interactive nature of the tool, which encourages repeated practice and reinforcement. However, without objective measures of improvement, these findings are tentative. Future studies should employ more rigorous methodologies, including control groups and objective assessments, to validate these preliminary observations.

This study offers initial insights, derived from student perceptions, into ChatGPT's potential role as an adjunctive tool in English language acquisition. Given the preliminary nature of these findings, there is a compelling case for extending the research to diverse university settings and varied student demographics. It would be pertinent to further explore the ideal frequency, duration, and efficacy of interactions with ChatGPT to better harness the advantages of chatbot-facilitated education. Additionally, the potential for chatbots to foster collaborative learning dynamics and enhance peer-to-peer support—especially within online and blended learning models—poses an intriguing direction for future research.

References

Ab Hamid, E. A. H., Maskur, H., & Mutalib, R. A. (2023). The Use of ChatGPT Applications in Learning: Impact on Understanding and Student Engagement in TVET Institutions. *Malaysian Journal of Information and Communication Technology (MyJICT)*, 78-87.

Abbas, N., Whitfield, J., Atwell, E., Bowman, H., Pickard, T., & Walker, A. (2022). Online chat and chatbots to enhance mature student engagement in higher education. *International Journal of Lifelong Education, 41*(3), 308-326. https://doi.org/10.1080/02601370.2022.2066213.

Akkara, S., Anumula, V., & Mallampalli, M. (2020). Impact of whatsapp interaction on improving L2 speaking skills. *International Journal of Emerging Technologies in Learning (iJET), 15*(3), 250-259. https://doi.org/10.3991/ijet.v15i03.11534.

Ali, Y. A., Ali, M. M. A., Yanbu, W. A., Abbas, S., & Hamid, S. (2025). The influence of ChatGPT on English grammar among ESL learners. *International Journal on Studies in Education*, *7*(3), 482-494.

Anantrasirichai, N., & Bull, D. (2022). Artificial intelligence in the creative industries: a review. *Artificial intelligence review, 55*(1), 589-656.

Ayedoun, E., Hayashi, Y., & Seta, K. (2015). A conversational agent to encourage willingness to communicate in the context of English as a foreign language. *Procedia Computer Science, 60*, 1433-1442. https://doi.org/10.1016/j.procs.2015.08.219.

Ayedoun, E., Hayashi, Y., & Seta, K. (2015). A conversational agent to encourage willingness to communicate in the context of English as a foreign language. *Procedia Computer Science, 60*, 1433-1442.

Bailin, S., Case, R., Coombs, J. R., & Daniels, L. B. (1999). Common misconceptions of critical thinking. *Journal of curriculum studies, 31*(3), 269-283. https://doi.org/10.1080/002202799183124.

Baskara, F. R. (2023). Integrating ChatGPT into EFL writing instruction: Benefits and challenges. *International Journal of Education and Learning, 5*(1), 44-55. https://doi.org/10.31763/ijele.v5i1.858.

Bower, M. (2017). *Design of Technology-enhanced Learning: Integrating Research and Practice*. United

Kingdom: Emerald Publishing.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology, 3*(2), 77-101. https://doi.org/10.1191/1478088706qp063oa.

Chen, J. A., Tutwiler, M. S., Metcalf, S. J., Kamarainen, A., Grotzer, T., & Dede, C. (2016). A multi-user virtual environment to support students' self-efficacy and interest in science: A latent growth model analysis. *Learning and Instruction, 41*, 11-22. https://doi.org/10.1016/j.learninstruc.2015.09.007.

Cotton, D. R., Cotton, P. A., & Shipway, J. R. (2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in education and teaching international, 61*(2), 228-239. https://doi.org/10.1080/14703297.2023.2190148.

Dashti, L., & Razmjoo, S. A. (2020). An examination of IELTS candidates' performances at different band scores of the speaking test: A quantitative and qualitative analysis. *Cogent Education, 7*(1), 1770936. https://doi.org/10.1080/2331186x.2020.1770936.

Demirdöken, G., & Okur, S. (2023). Psychometric properties of speaking anxiety scale and an interdisciplinary investigation with serial mediation analysis. *Innovation in Language Learning and Teaching, 17*(3), 706-722. https://doi.org/10.1080/17501229.2022.2123920.

Facione, P. A. (2000). The disposition toward critical thinking: Its character, measurement, and relationship to critical thinking skill. *Informal logic, 20*(1). https://doi.org/10.22329/il.v20i1.2254.

Fryer, L. K., Ainley, M., Thompson, A., Gibson, A., & Sherlock, Z. (2017). Stimulating and sustaining interest in a language course: An experimental comparison of Chatbot and Human task partners. *Computers in human behavior, 75*, 461-468. https://doi.org/10.1016/j.chb.2017.05.045.

Fryer, L. K., Nakao, K., & Thompson, A. (2019). Chatbot learning partners: Connecting learning experiences, interest and competence. *Computers in human Behavior, 93*, 279-289. https://doi.org/10.1016/j.chb.2018.12.023.

Fryer, L., Coniam, D., Carpenter, R., & Lăpușneanu, D. (2020). *Bots for Language Learning Now: Current and Future Directions.* Handle Proxy. http://hdl.handle.net/10125/44719.

Gallacher, A., Thompson, A., Howarth, M., Taalas, P., Jalkanen, J., Bradley, L., & Thouësny, S. (2018). "My robot is an idiot!"–Students' perceptions of AI in the L2 classroom. *Future-proof CALL: language learning as exploration and encounters–short papers from EUROCALL,* 70-76. https://doi.org/10.14705/rpnet.2018.26.815.

García-Peñalvo, F. J. (2023). La percepción de la Inteligencia Artificial en contextos educativos tras el lanzamiento de ChatGPT: disrupción o pánico. *Education in the Knowledge Society (EKS), 24*, e31279-e31279. https://doi.org/10.14201/eks.31279.

Goar, V., Yadav, N. S., & Yadav, P. S. (2023). Conversational AI for natural language processing: An review of ChatGPT. *International Journal on Recent and Innovation Trends in Computing and Communication, 11*, 109-117. https://doi.org/10.17762/ijritcc.v11i3s.6161.

Gray, L. (2014). Book Review: The American Heritage Dictionary of the English Language. *Journal of English Linguistics, 42*(2), 176-180.

Gunawardena, C. N. (1995). Social presence theory and implications for interaction and collaborative learning in computer conferences. *International journal of educational telecommunications, 1*(2), 147-166. https://www.learntechlib.org/primary/p/15156/.

Halpern, D. F. (1998). Teaching Critical Thinking for Transfer across Domains: Disposition, Skills, Structure Training, and Metacognitive Monitoring. *American Psychologist, 53*(4), 449–55. https://doi.org/10.1037/0003-066x.53.4.449.

Hamzah, M. I., L. Li, & Sani, S. S. A. (2021) The Potential of Chatbots in Facilitating Learning in Higher Education: A Preliminary Study. *International Journal of Emerging Technologies in Learning (iJET), 16*(01), 4–20.

Han, J., Yoo, H., Myung, J., Kim, M., Lee, T. Y., Ahn, S. Y., & Oh, A. (2023). ChEDDAR: Student-ChatGPT dialogue in EFL writing education. *arxiv preprint arxiv:2309.13243*. https://arxiv.org/abs/2309.13243.

Han, M., & Niu, S. (2021). Application of virtual scenario teaching in spoken English teaching. *International Journal of Emerging Technologies in Learning (iJET), 16*(18), 129-142. https://www.learntechlib.org/p/220125/.

Han, X., & Huwan, T. (2022). [Retracted] The Modular Design of an English Pronunciation Level Evaluation System Based on Machine Learning. *Security and Communication Networks, 2022*(1), 6804131. https://doi.org/10.1155/2022/6804131.

Haristiani, N., & Rifai, M. M. (2021). Chatbot-based application development and implementation as an autonomous language learning medium. *Indonesian Journal of Science and Technology, 6*(3), 561-576. https://doi.org/10.17509/ijost.v6i3.39150.

Harris, J., Golinkoff, R. M., & Hirsh-Pasek, K. (2011). Lessons from the crib for the classroom: How children really learn vocabulary. *Handbook of early literacy research, 3*, 49-65.

Hong, J. C., Hwang, M. Y., Liu, M. C., Ho, H. Y., & Chen, Y. L. (2014). Using a "prediction–observation–explanation" inquiry model to enhance student interest and intention to continue science learning predicted by their Internet cognitive failure. *Computers & Education, 72*, 110-120. https://doi.org/10.1016/j.compedu.2013.10.004.

Hsu, L. (2022). To CALL or not to CALL: empirical evidence from neuroscience. *Computer Assisted Language Learning, 35*(4), 792-815. https://doi.org/10.1080/09588221.2020.1750429.

Hwang, J. K., Mancilla-Martinez, J., McClain, J. B., Oh, M. H., & Flores, I. (2020). Spanish-speaking English learners' English language and literacy skills: The predictive role of conceptually scored vocabulary. *Applied Psycholinguistics, 41*(1), 1-24. https://doi.org/10.1017/s0142716419000365.

Iio, T., Maeda, R., Ogawa, K., Yoshikawa, Y., Ishiguro, H., Suzuki, K., ... & Hama, M. (2019). Improvement of Japanese adults' English speaking skills via experiences speaking to a robot. *Journal of Computer Assisted Learning, 35*(2), 228-245.

Javaid, M., Haleem, A., Singh, R. P., Khan, S., & Khan, I. H. (2023). Unlocking the opportunities through ChatGPT Tool towards ameliorating the education system. *BenchCouncil Transactions on Benchmarks, Standards and Evaluations, 3*(2), 100115. https://doi.org/10.1016/j.tbench.2023.100115.

Jeon, J. (2023). Chatbot-assisted dynamic assessment (CA-DA) for L2 vocabulary learning and diagnosis. *Computer Assisted Language Learning,36*(7), 1338-1364. https://doi.org/10.1080/09588221.2021.1987272.

Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., ... & Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and individual differences, 103*, 102274. https://doi.org/10.1016/j.lindif.2023.102274.

Kim, H. S., & Kim, N. Y. (2021). Effects of AI Chatbots on EFL Students' Communication Skills. *Korean Journal of English Language and Linguistics, 21*, 712-734. https://doi.org/ 10.15738/kjell.21.202108.712.

King, M. R., & ChatGPT. (2023). A conversation on artificial intelligence, chatbots, and plagiarism in higher education. *Cellular and molecular bioengineering, 16*(1), 1-2.

Kohnke, L. (2023). L2 learners' perceptions of a chatbot as a potential independent language learning tool. *International Journal of Mobile Learning and Organisation, 17*(1-2), 214-226. https://doi.org/10.1504/ijmlo.2023.128339.

Kürüm, E. Y. (2016). Teaching speaking skills. *Annals of the University of Oradea: Economic Science, 25*(1), 264-273.

Limo, F. A. F., Tiza, D. R. H., Roque, M. M., Herrera, E. E., Murillo, J. P. M., Huallpa, J. J., ... & Gonzáles, J. L. A. (2023). Personalized tutoring: ChatGPT as a virtual tutor for personalized learning experiences. *Przestrzeń Społeczna (Social Space), 23*(1), 293-312..

Lin, M. P. C., & Chang, D. (2020). Enhancing post-secondary writers' writing skills with a chatbot. *Journal of Educational Technology & Society, 23*(1), 78-92. https://www.jstor.org/stable/26915408.

Liu, Y. (2021). Integrating Various Resources In Supporting The Learning Of English Learners. *Kotesol Proceedings 2021*, 231.

Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. *Education Sciences, 13*(4), 410. https://doi.org/10.3390/educsci13040410.

Mahmoud, R. H. (2022). Implementing AI-Based Conversational Chatbots in EFL Speaking Classes: An Evolutionary Perspective, *Research Square*, 1-21. https://doi.org/10.21203/rs.3.rs-1911791/v1.

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.

Naumova, E. N. (2023). A mistake-find exercise: a teacher's tool to engage with information innovations, ChatGPT, and their analogs. *Journal of Public Health Policy, 44*(2), 173-178. https://doi.org/10.1057/s41271-023-00400-1.

Pan, D., & Zhou, H. (2018). English Learning System Design for College Students Personalized English Grammar Check and Diagnosis. *International Journal of Emerging Technologies in Learning, 13*(4). https://doi.org/10.3991/ijet.v13i04.8467.

Pardos, Z. A., & Bhandari, S. (2023). Learning gain differences between ChatGPT and human tutor generated algebra hints. *arxiv preprint arxiv:2302.06871*. https://arxiv.org/abs/2302.06871.

Rudolph, J., Tan, S., & Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education?. *Journal of applied learning and teaching, 6*(1), 342-363. https://doi.org/10.37074/jalt.2023.6.1.9.

Sánchez, O. V. G. (2023). Uso y percepción de ChatGPT en la educación superior. *Revista de Investigación en Tecnologías de la Información, 11*(23), 98-107. https://doi.org/10.36825/riti.11.23.009.

Sánchez-Ruiz, L. M., Moll-López, S., Nuñez-Pérez, A., Moraño-Fernández, J. A., & Vega-Fleitas, E. (2023). ChatGPT challenges blended learning methodologies in engineering education: a case study in mathematics. *Applied Sciences, 13*(10), 6039. https://doi.org/10.3390/app13106039

Santhanasamy, C., & Yunus, M. M. (2022). The Flipped Learning and Blendspace to Improve Pupils' Speaking

Skills. *Frontiers in Psychology, 13*, 866270 . https://doi.org/10.3389/fpsyg.2022.866270.

Short, J., Williams, E., & Christie, B. (1976). The Social Psychology of Telecommunications. *American Sociological Association, 7*(1), 32-33.

Shoufan, A. (2023). Exploring students' perceptions of ChatGPT: Thematic analysis and follow-up survey. *IEEE Access, 11*, 38805-38818. https://doi.org/10.1109/ACCESS.2023.3268224.

Tegos, S., Demetriadis, S., & Tsiatsos, T. (2014). A configurable conversational agent to trigger students' productive dialogue: a pilot study in the CALL domain. *International Journal of Artificial Intelligence in Education, 24*, 62-91.

Thorp, H. H. (2023). ChatGPT is fun, but not an author. *Science, 379*(6630), 313-313. https://doi.org/10.1126/science.adg7879.

Tsai, S. C. (2015). Implementing courseware as the primary mode of task-based ESP instruction: A case study of EFL students. *Computer Assisted Language Learning, 28*(2), 171-186. https://doi.org/10.1080/09588221.2013.818554.

Ulla, M. B., Perales, W. F., & Busbus, S. O. (2023). 'To generate or stop generating response': Exploring EFL teachers' perspectives on ChatGPT in English language teaching in Thailand. *Learning: Research and Practice, 9*(2), 168-182.

Wahyuni, D. S. (2022). Integrated Classroom-Chatbot Experience: an Alternative Solution for English as Foreign Language Learners. *English Language Education and Current Trends (ELECT)*, 63-68. https://doi.org/10.37301/elect.v1i1.36

Wang, X., Liu, Q., Pang, H., Tan, S. C., Lei, J., Wallace, M. P., & Li, L. (2023). What matters in AI-supported learning: A study of human-AI interactions in language learning using cluster analysis and epistemic network analysis. *Computers & Education, 194*, 104703. https://doi.org/10.1016/j.compedu.2022.104703.

Wang, Y. F., Petrina, S., & Feng, F. (2017). VILLAGE—V irtual I mmersive L anguage L earning and G aming E nvironment: Immersion and presence. *British Journal of Educational Technology, 48*(2), 431-450.

West, J. P. (2015). I Still Can't Speak English: Make Your Own Free Social Learning, Real Practice English Course and Finally Speak English Comfortably. Languages Out There Ltd.

Xiao, W. Q., & Park, M. Y. (2021). Using Automatic Speech Recognition to Facilitate English Pronunciation Assessment and Learning in an EFL Context. *International Journal of Computer-Assisted Language Learning and Teaching, 11*(3), 74–91. https://doi.org/10.4018/ijcallt.2021070105.

Xiao, Y. Y., & Zhi, Y. Y. (2023). An Exploratory Study of EFL Learners' Use of ChatGPT for Language Learning Tasks: Experience and Perceptions. *Languages, 8*(3), 212. https:doi.org/10.3390/langauges8030212.

Xu, Y. (2022). English speech recognition and evaluation of pronunciation quality using deep learning. *Mobile Information Systems, 2022*(1), 7186375. https://doi.org/10.1155/2022/7186375.

Xu, Y., & Warschauer, M. (2020, June). Exploring young children's engagement in joint reading with a conversational agent. In *Proceedings of the interaction design and children conference* (pp. 216-228). https://doi.org/10.1145/3392063.3394417.

Zaghlool, Z. D., & Khasawneh, M. A. S. (2023). Incorporating the Impacts and Limitations of AI-Driven Feedback, Evaluation, and Real-Time Conversation Tools in Foreign Language Learning. *Migration Letters, 20*(7), 1071-1083.

Author Information

Huanhuan Zhang

https://orcid.org/0009-0008-6565-1118

Macao Polytechnic University

R. de Luís Gonzaga Gomes, Macao Polytechnique University

Macao, China

Yujie Su

https://orcid.org/0000-0003-1444-1598

Wenzhou University

Wenzhou, Zhejiang

China

XiaoShu Xu (Corresponding Author)

https://orcid.org/0000-0002-0667-4511

Wenzhou University

Wenzhou, Zhejiang

China

Contact e-mail: *Lisaxu@wzu.edu.cn*

Vivian Ngan-Lin Lei

https://orcid.org/0000-0002-9995-8432

Macao Polytechnic University

R. de Luís Gonzaga Gomes, Macao Polytechnique University

Macao, China

Shanshan Hao

https://orcid.org/0009-0000-8460-5653

Macao Polytechnic University

R. de Luís Gonzaga Gomes, Macao Polytechnique University

Macao, China