



AI-Supported Language Learning for Developing English Listening and Speaking Skills and ChatGPT Literacy*

Hea-Suk Kim (Seoul Women's University)



This is an open-access article distributed under the terms of the Creative Commons License, which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Received: August 19, 2025

Revised: September 11, 2025

Accepted: October 10, 2025

Kim, Hea-Suk
Associate Professor, College of
General Education
Seoul Women's University
Tel: 02-970-7862
Email: shskim@swu.ac.kr

* This work was supported by a research grant from Seoul Women's University (2025-0132).

ABSTRACT

Kim, Hea-Suk. 2025. AI-Supported language learning for developing English listening and speaking skills and ChatGPT literacy. *Korean Journal of English Language and Linguistics* 25, 1353-1377.

This study explored the impact of AI-supported instruction on the development of English communicative skills and ChatGPT literacy among Korean university students in an EFL context. In the spring semester of 2024, the experiment was conducted with students enrolled in a college English course and lasted for one semester. Participants were divided into a control group (N = 25), who practiced speaking through face-to-face conversations, and an experimental group (N = 26), who engaged in practice using ChatGPT. Data were collected through pre- and post-tests in speaking and listening, as well as surveys assessing ChatGPT literacy. While both groups showed significant within-group improvement in speaking and listening skills, no statistically significant differences between them. However, learners in the ChatGPT group demonstrated notably higher gains in speaking performance. The experimental group showed growth across five domains of ChatGPT literacy: technical proficiency, critical evaluation, communication proficiency, creative application, and ethical competence. Learners reported overall positive experiences, citing improved speaking skill, increased confidence, convenience, and engagement, while also noting limitations such as technical issues and occasional inaccuracies. These findings offer deeper insights into AI-integrated language instruction and highlight the importance of prompt literacy and pedagogical support in fostering autonomous learning in digital contexts.

KEYWORDS

Generative AI, ChatGPT, speaking, listening, ChatGPT literacy, perception

1. Introduction

Modern society is undergoing rapid transformation driven by digitalization, which continues to serve as a pivotal force for global development. Within this digital shift, a vast information ecosystem is emerging, accompanied by increasing demand for advanced technologies capable of efficiently managing and analyzing data. As digital systems evolve at an accelerated pace, the integration of artificial intelligence (AI) has also gained momentum. In particular, significant advancements in natural language processing (NLP) and AI technologies are exerting tangible influence across various sectors. As these tools become more user-friendly, the general public has gained broader access to their applications.

With the spread of AI, digital literacy has evolved into AI literacy, encompassing technical skills, an understanding of algorithms and data, critical evaluation, and ethical awareness (Cox 2023, Dignum 2019, Kim and Kim 2024). In education, AI literacy has been shown to improve learners' self-efficacy and academic performance (Yim and Su 2024). ChatGPT literacy, a subfield of AI literacy, emphasizes practical use of generative AI tools. In language education, it highlights technical proficiency, critical thinking, communication, creativity, and ethics (Lee and Park 2024, Liu 2025, Ma et al. 2024).

Recent research has explored diverse aspects of AI use in education, including learner perceptions (Bao and Li 2023, Choe 2023, Choi and Choe 2024, Im 2023, Lee and Lee 2021, Oh 2023, Shin and Lee, 2024), teacher preparation and professional development (Jeon and Lee 2023), feedback mechanisms (Guo and Wang 2023, Young and Shishido 2023), linguistic competence (Baskara and Mukarto 2023, Jeong 2024, Kim et al. 2025), and AI or prompt literacy (Choi and Choe 2024, Hwang 2023, Yang 2025). These lines of inquiry reflect a broader pedagogical shift toward exploring the educational potential of generative AI tools.

Among these tools, OpenAI's ChatGPT has attracted particular attention for its ability to perform complex, creative, and intellectually engaging tasks, demonstrating the transformative potential of digital technology in education. ChatGPT has emerged as a valuable instructional resource across various curricular areas, particularly in foreign language learning (Baskara and Mukarto 2023, Kim and Kim 2024, Kohnke et al. 2023, Sharples 2023), and its integration into classrooms is reshaping pedagogical practices.

Although many studies report positive outcomes—especially in writing tasks—based on learners' perceptions of using ChatGPT in foreign language learning (Choe 2023, Liu et al. 2024, Shim and Kim 2024), concerns remain. Some educators caution that excessive reliance on ChatGPT may hinder the development of learners' self-directed learning skills (Jeong 2024), while others question its ability to support holistic language acquisition (Cha 2023). These critiques highlight the importance of strategic AI integration—one that maximizes learning outcomes, prevents misuse, and promotes balanced development. Notably, among the four core language competencies—reading, writing, speaking, and listening—existing research has overwhelmingly concentrated on writing, with particular emphasis on how learners engage with AI tools to enhance their written output (Im 2023, Jeong 2024, Oh 2023). While learner perceptions have also received attention (Choe 2023, Liu et al. 2024), this dominant focus on writing has led to a significant gap in empirical studies on AI-supported speaking and listening practices, leaving these essential communicative domains comparatively underexplored. In English as a Foreign Language (EFL) contexts, where opportunities for authentic spoken interaction are often limited, technological approaches that support learners' communicative skills are especially critical.

Furthermore, there is a striking lack of research on ChatGPT literacy—a key digital competency involving the critical understanding, purposeful use, and ethical engagement with generative AI tools (Lee and Park 2023). Developing this literacy goes beyond basic tool usage; it empowers learners to interact meaningfully and responsibly with AI in educational contexts.

To address these gaps, the present study aims to examine the impact of ChatGPT-based English conversation practice on Korean university students' speaking and listening skills. Additionally, it investigates how learners' ChatGPT literacy evolves through sustained interaction over the course of a semester. Rather than merely evaluating the tool's functional efficacy, this research seeks to provide empirical evidence for new pedagogical directions and instructional strategies in AI-mediated language education. Accordingly, this study addresses the following research questions:

1. To what extent does AI-supported speaking practice enhance learners' English speaking and listening skills in an EFL context?
2. How does engaging in AI-supported language learning contribute to the development of ChatGPT literacy among university students?
3. What are learners' perceptions and experiences of using ChatGPT for speaking practice over the course of a semester?

2. Literature Review

2.1 Digital Literacy, AI Literacy and ChatGPT Literacy

Since the onset of the digital revolution, digital literacy plays a pivotal role in acquiring information, making informed decisions, and solving problems—functions that directly influence both learning outcomes and overall quality of life. This literacy has not only reshaped societal structures but has also significantly enhanced individual competencies across diverse domains (Raschke 2003). Individuals lacking digital competence risk falling into significant knowledge gaps, underscoring the urgency of cultivating these skills. Gilster (1997), in one of the earliest conceptualizations of digital literacy, broadly defined it as the ability to understand and utilize information derived from digital sources. This definition now encompasses interactions with AI-powered tools such as ChatGPT. Additionally, Gilster emphasized the need to foster adaptable thinkers equipped not only with technical skills but also with intangible competencies—such as a perspective aligned with the digital era and a willingness to engage with interactive media. This foundational perspective established a framework for understanding human engagement with emerging technologies. As digital landscapes continue to evolve, the notion of literacy must likewise expand, particularly in light of the increasing prevalence of artificial intelligence. Within this expanded scope, AI literacy emerges as a vital extension of digital literacy, emphasizing not only the ability to operate AI tools but also the capacity to understand their underlying processes and evaluate their broader societal implications.

AI literacy is increasingly recognized as a multidimensional competency essential for navigating the digital age. It encompasses not only the ability to use AI technologies but also a foundational understanding of how these systems operate, the capacity to apply them in diverse contexts, and the awareness of their broader social and ethical implications (Kim and Kim 2024). As Cox (2023) outlines, AI literacy consists of four key domains: (1) conceptual understanding of core principles such as machine learning, algorithms, and data processing; (2) practical application skills that enable individuals to use AI tools for problem-solving and decision-making; (3) critical evaluation abilities to assess the accuracy, fairness, and transparency of AI systems; and (4) ethical awareness regarding issues like privacy, bias, and the impact of automation on employment (Dignum 2019). From an educational standpoint, fostering AI literacy has been shown to enhance learners' self-efficacy, promote positive emotional engagement with AI-integrated learning environments, and improve academic performance (Yim and

Su 2024). These findings underscore the growing necessity for educational frameworks that equip individuals with the skills and mindset to responsibly engage with AI technologies. As AI becomes increasingly embedded across disciplines and industries, cultivating AI literacy is not merely a technical requirement but a foundational element of modern education. Reflecting this trend, the integration of generative AI-based chatbots—most notably ChatGPT developed by OpenAI—into various educational contexts has become increasingly prevalent, offering tangible opportunities to cultivate the competencies encompassed within AI literacy.

AI literacy equips educators and technology specialists with the competencies needed to systematically integrate AI-based tools into educational curricula. In contrast, ChatGPT literacy supports language teachers in utilizing ChatGPT to develop effective instructional materials and deliver personalized guidance tailored to learners' levels and needs (Liu 2025). For language learners, ChatGPT literacy serves as a practical skill set that enables them to explore adaptive learning technologies, receive customized feedback, and actively engage with ChatGPT resources. While in-depth technical expertise may enhance comprehension, effective integration of the tool into language education primarily requires a practical understanding grounded in instructional experience. Ultimately, AI literacy provides a broader framework that encompasses ChatGPT literacy, including the specific skills and knowledge necessary for its educational application (Liu 2025). By aligning these two literacies, educators and learners can confidently navigate the evolving landscape of AI technologies and ethically maximize the educational potential of ChatGPT (Ma et al. 2024).

As ChatGPT's influence on society continues to expand, the ability to use it effectively is increasingly recognized as a vital life skill (Lee and Park 2023). Lee and Park investigated the relationship between ChatGPT proficiency and user satisfaction, emphasizing the mediating roles of two key user motivations: acquiring information and knowledge, and engaging in entertainment and leisure. The study surveyed 822 university students with prior experience using ChatGPT. Results indicated that higher levels of ChatGPT literacy were positively correlated with greater user satisfaction, and that the motivations for learning and enjoyment partially mediated this relationship. Building on these findings, Lee and Park conducted a follow-up study in 2024 to conceptualize ChatGPT literacy more rigorously. Through a comprehensive review of existing literature and expert focus group discussions, they identified five core components of ChatGPT literacy: technical proficiency, critical evaluation, communication proficiency, creative application, and ethical competence.

Another study (Ma et al. 2024) proposed a six-component framework—benefits, limitations, prompts, evaluation, assessment, and ethics—highlighting the competencies language teachers need to use ChatGPT effectively. Similarly, Lund et al. (2023) identified three related literacies: information, data, and privacy literacy. While ChatGPT use was found to enhance information and privacy literacy, it had limited impact on data literacy.

In conclusion, both AI literacy and ChatGPT literacy represent essential competencies for navigating contemporary educational environments, sharing a multidimensional nature that encompasses technical proficiency, ethical awareness, and critical thinking. AI literacy offers a comprehensive framework that includes conceptual understanding of machine learning and algorithms, practical application across diverse contexts, and the ability to critically evaluate and ethically engage with AI systems. In contrast, ChatGPT literacy functions as a more specialized subset of AI literacy, focusing on the practical use of generative AI tools—particularly in language education. Language teachers and learners can employ ChatGPT to create personalized instructional materials and adaptive learning experiences, even without fully grasping the underlying technical architecture. Ultimately, aligning and integrating these two literacies empowers educators and learners to confidently navigate the evolving landscape of AI technologies and to harness the educational potential of ChatGPT in a responsible and ethical manner.

In light of this perspective, the present study foregrounds the concept of ChatGPT literacy, with a particular

focus on how learners interact with this widely adopted AI tool in foreign language learning contexts. Rather than treating ChatGPT merely as a technological artifact, the study conceptualizes learner engagement as central to understanding ChatGPT literacy. Drawing on the framework proposed by Lee and Park (2024), ChatGPT literacy is delineated into five interrelated components: technical proficiency, critical evaluation, communication proficiency, creative application, and ethical competence. These dimensions collectively offer a comprehensive lens through which to examine learners' evolving relationship with AI-driven language tools.

2.2 AI-supported Language Learning

The advent of the Fourth Industrial Revolution—driven by rapid advancements in science and technology—has positioned intelligent information systems and artificial intelligence (AI) as transformative forces on a global scale. Among these innovations, recent progress in natural language processing (NLP) has notably amplified AI's impact, extending its reach across various domains, including education. As AI tools become more accessible and user-friendly, their integration into diverse educational contexts has accelerated, offering new possibilities for teaching, learning, and communication.

In the field of education, the rise of generative AI has introduced innovative instructional tools, with ChatGPT, standing out as a particularly impactful resource for foreign language teaching and learning (Baskara and Mukarto 2023, Kohnke et al. 2023, Sharples 2023). As a large language model (LLM) built on advanced transformer-based pre-training, ChatGPT distinguishes itself from earlier AI tools through its versatility, depth of understanding, and creative output. Its ability to generate coherent and contextually appropriate responses to a wide range of prompts has opened new possibilities for language acquisition.

From a theoretical perspective, the use of ChatGPT aligns with the second language acquisition (SLA) model proposed by Gass et al. (2020), which conceptualizes language learning as a complex cognitive and social process. This model emphasizes how learners perceive and process input, engage in interaction, and produce output, with feedback playing a central role in language development. Gass highlights interaction and feedback as critical components, arguing that learners improve through active communication, error correction, and repeated use of language in meaningful contexts. In this regard, ChatGPT offers rich input, real-time feedback, and interactive dialogue, creating a balanced SLA (Second Language Acquisition) environment that integrates both cognitive and social dimensions. For example, it supports vocabulary and grammar acquisition (Baskara and Mukarto 2023, Kim and Kim 2024), enhances comprehension through translation (Kim et al. 2025) and explanation features, facilitates integration through follow-up questions (Kohnke et al. 2023), and reinforces learning through personalized practice activities.

Pedagogically, ChatGPT enables teachers to design individualized learning environments (Baidoo-Anu and Ansah 2023), and generate texts and dialogues (Crosthwaite and Baisa 2023, Kim and Kim 2024). It also automates tasks such as essay writing, grading, and feedback provision (Jeong 2024, Mizumoto and Eguchi 2023), thereby fostering self-directed language learning (Lee and Lee 2021). Moreover, ChatGPT-based instruction has been shown to enhance learners' motivation, sustaining interest and engagement in language learning (Kim and Kim 2024).

Recent studies have increasingly focused on teacher and student attitudes, intentions, and perceptions regarding the use of ChatGPT for language learning (Bao and Li 2023, Choi and Choe 2024, Im 2023, Synekop et al. 2024). Research by Liu and Ma (2024) emphasizes the importance of learners' perceptions—such as perceived usefulness and satisfaction—in shaping their attitudes and intentions toward ChatGPT use (Choi and Choe 2024, Kim and Kim 2024).

Synekop et al. (2024) conducted a study at a Ukrainian university involving 22 teachers and 60 students, examining attitudes toward the use of ChatGPT in English classes. The study focused on short conversation activities and explored how learners engaged with ChatGPT from the perspective of academic integrity, ultimately reporting that such interactions positively influenced students' English communication skills. Quantitative survey results revealed that students held positive attitudes, while teachers adopted a more neutral stance. Both groups agreed on ChatGPT's potential to enhance communicative competence and emphasized the need to integrate AI tools into ESP courses to support language learning and adapt to evolving educational contexts.

Scholars have explored its effectiveness in generating educational content and providing feedback in EFL contexts (Guo and Wang 2023, Young and Shishido 2023), as well as in automatic question generation and telecollaboration (McCallum 2023). Other studies highlight ChatGPT's role in enhancing learner engagement and concentration, thereby improving language learning outcomes (Im 2023, Kim et al. 2025, Kim and Kim 2024, R. Kim 2024, Lee and Lee 2021). Lee and Lee (2021) suggest that students' emotional experiences in English learning are shaped by their engagement in AI environments and sociocultural motivational structures. To foster effective second language (L2) acquisition, educators should support learners in developing self-directed learning and promote enjoyment throughout the learning process.

Research has also explored ChatGPT's potential across the four language skills, with writing receiving the most attention (Choe 2023, Im 2023, Jeong 2024, Kim et al. 2023, Liu et al. 2024, Oh 2023, Shim and Kim 2024). For instance, ChatGPT facilitates interactive communication and provides students with personalized feedback on their writing (Jeong 2024). Choe (2023) conducted a qualitative study on the impact of ChatGPT-assisted summary writing activities on Korean pre-service English teachers. The findings indicated improvements in vocabulary choice, text organization, and awareness of summary quality, supported by individualized and continuous feedback. These educational and technical experiences also had positive emotional effects, reducing anxiety and increasing confidence and willingness to use ChatGPT. Similarly, Oh (2023) examined the use of ChatGPT for essay writing following thematic reading activities. Survey results showed that students viewed ChatGPT as an efficient tool for idea generation, information gathering, and time management. However, concerns about information quality and reliability highlighted the need for instructional support, including strategies for source comparison, critical evaluation, and clear ethical guidelines.

On the other hand, Cheon (2023) analyzed the impact of using ChatGPT, in general English courses on EFL university students' speaking proficiency. A four-week experiment was conducted with 80 students at a university in Korea, where the experimental group engaged in situation-based conversational practice using ChatGPT in travel-related contexts. Based on data collected through pre- and post-speaking assessments, surveys, and in-depth interviews, the experimental group showed greater improvement than the control group, particularly in vocabulary and cultural appropriateness. Participants highlighted several advantages of using ChatGPT as a learning aid, including interactive Q&A in conversational format, ease of use with simple text input, level-specific personalized learning, enhanced concentration, and reduced anxiety related to speaking English. Another study examined the effects of ChatGPT-based learning on Korean university students' English speaking proficiency in an EFL context (Noh 2024). Comparing an experimental group using ChatGPT with a control group receiving traditional instruction, the research assessed vocabulary, pronunciation, grammar, and fluency through pre- and post-tests, weekly tasks, surveys, and interviews. Findings showed that ChatGPT use significantly improved vocabulary, pronunciation, and fluency, while lowering psychological barriers to speaking.

In reading instruction, Kim et al. (2025) integrated ChatGPT into EFL reading classes and reported significant improvements across five literacy domains: technical proficiency, critical evaluation, communication, creative application, and ethical awareness. While students viewed ChatGPT as an effective tool for achieving academic

goals, they maintained a cautious stance regarding information accuracy and ethical use. Interview data revealed high trust in ChatGPT, tempered by strategies such as follow-up questioning, refinement of prompts, and cross-verification with reliable sources. Ethical concerns, including plagiarism prevention and responsible use of AI-generated content, were also emphasized.

Other studies have investigated usage prompt patterns (Huang 2024, Lee 2024, Yang 2025). Lee (2024) explored ChatGPT as a feedback tool in English writing classes, analyzing prompt usage among 31 pre-service teachers grouped by proficiency level. Higher-level students used prompts more frequently and engaged in sustained, interactive dialogue. Learners expressed satisfaction with ChatGPT's feedback on vocabulary, grammar, and sentence construction. Huang (2024) similarly found that interaction with ChatGPT facilitated dynamic practice and personalized feedback, contributing to speaking skill development through well-structured prompts that elicited meaningful responses.

Although various previous studies have reported that generative AI tools such as ChatGPT have a positive impact on foreign language learning, some scholars caution that these tools may have adverse effects. They warn that reliance on AI could lead to diminished language production skills, reduced critical thinking, ethical concerns and increased learner dependency (Bok and Cho 2023, Kim et al. 2025, Yang 2025).

In the study of Kim et al. (2024), most experimental studies in Korea focused on participants' perceptions of ChatGPT, with 71.4% falling into this category. Only 9.5% studied both reading and writing, and 4.8% each covered speaking and vocabulary. No studies addressed listening or grammar. Internationally, 77.3% of studies also focused on user perceptions. In addition, international research on ChatGPT has primarily focused on its application across different language skills, with a strong emphasis on writing, feedback provision, and the enhancement of learners' motivation and engagement. Most of these studies have employed survey-based perception research, followed by small-scale experiments and case studies.

Taken together, these studies suggest that ChatGPT serves as an innovative tool in foreign language education, supporting both emotional and technical aspects of learning. However, further research is needed—particularly in the domain of speaking and listening—to fully realize its potential and ensure its effective integration into language instruction.

3. Method

3.1 Participants

This study was conducted during the first semester of 2024, targeting first-year students enrolled in the “College English: Listening and Speaking” course. The course is a required subject for freshmen and carries two academic credits. Prior to enrollment, freshmen were administered a mock TOEIC test by the university. Based on the results, students were classified into high, intermediate, and low proficiency levels, and each student selected a class corresponding to their assigned level. The participants in this study belonged to the low proficiency group, with an average TOEIC score of approximately 310 (Control Group = 310.15; Experimental Group = 313.53). The course was designed specifically for first-year students and offered a structured curriculum aimed at improving listening and speaking skills. A total of 51 students participated in the experiment, with 25 assigned to the control group and 26 to the experimental group. These students were placed into level-specific classes through the university's diagnostic process and voluntarily registered for the course section appropriate to their proficiency level.

Table 1. Demographic Information about Control vs. Experimental Groups

Group	Age	Major	Proficiency (CSAT)	Mock TOEIC	ChatGPT Use	ChatGPT Use for English
Control (N=25)	19-23	Natural Sciences (11)	Second-grade (4)	310.15	Yes (10)	Yes (1)
		Arts & Physical Education (9)	Third- grade (5)		No (15)	No (24)
		Future Industrial Convergence (5)	Fourth-grade (5)			
			Fifth-grade/below (11)			
Experimental (N=26)	19-24	Arts & Physical Education (10)	Second-grade (2)	313.53	Yes (5)	Yes (2)
		Future Industrial Convergence (11)	Third- grade (5)		No (21)	No (24)
		Natural Sciences (5)	Fourth-grade (8)			
			Fifth-grade/below (11)			

Note: (Number of the students)

As seen in Table 1, participants in the study were divided into two groups: the control group (N = 25) and the experimental group (N = 26). The age range for both groups was relatively similar, from 19 to 23 in the control group and 19 to 24 in the experimental group. In terms of majors, the control group included students from Natural Sciences (11), Arts and Physical Education (9), and Future Industrial Convergence (5). The experimental group showed a slightly different distribution with Arts and Physical Education (10), Future Industrial Convergence (11), and Natural Sciences (5). English proficiency, as measured by CSAT (College Scholastic Ability Test) scores, varied across groups. In the control group, 14 students placed in second to fourth grade levels (Second: 4, Third: 5, Fourth: 5), while 11 students scored below fifth grade. The experimental group showed a comparable pattern with Second (2), Third (5), Fourth (8), and 11 students below fifth grade.

Regarding prior exposure to ChatGPT, ten students in the control group and five students in the experimental group had previously used the tool. However, when asked specifically about using ChatGPT for English learning, only one student in the control group and two students in the experimental group reported such experience. These findings reflect a generally low level of prior engagement with ChatGPT for educational purposes across both groups, underscoring the importance of investigating its potential instructional impact in a structured learning context.

Table 2. Overview of Skills in Control and Experimental Groups

	Group	Listening	Speaking	Reading	Writing	Grammar	Vocabulary
Present: confident skills	Cont.	15	2	13	5	2	1
	Exp.	15	1	10	3	3	1
Future: skills to develop	Cont.	15	25	3	4	5	6
	Exp.	17	22	4	4	4	5
ChatGPT use for learning	Cont.	4	14	3	10	7	6
	Exp.	5	15	3	12	5	5

Note: Cont. Control Group; Exp. Experimental Group

Table 2 summarizes students' self-perceptions regarding various language learning domains, highlighting (1) areas where they currently feel confident, (2) areas they wish to improve, and (3) areas where they expect to grow through the use of ChatGPT tools. Each skill domain is presented with corresponding frequencies.

The domains students felt most confident in were Listening (15 responses) and Reading (13 responses), suggesting a higher level of self-assurance in receptive language skills. This indicates students are generally

comfortable with comprehension-based and input-driven learning activities.

In contrast, speaking (25 responses) overwhelmingly appeared as the area students most wished to improve. This reflects a strong desire to enhance expressive abilities and a keen motivation toward more active language production. Vocabulary, Grammar, and Writing also emerged as notable areas for growth, indicating a student-driven interest in improving language accuracy and richness.

Regarding areas where students expect ChatGPT tools to support their development, speaking (14 responses) and Writing (10 responses) were most frequently mentioned. This suggests that generative AI tools are perceived as valuable aids for expressive learning. Furthermore, grammar and vocabulary were also cited as domains where students anticipated receiving more effective feedback and learning materials through ChatGPT. Overall, the data suggest that students demonstrated high motivation and a clear desire to improve their speaking proficiency. This finding reinforces the relevance of the study's focus on speaking enhancement through ChatGPT-supported instruction and highlights its potential in addressing learners' linguistic needs.

3.2 Teaching Procedures

This study was conducted within a mandatory freshman-level course titled College English: Listening and speaking, which ran for a total of 15 weeks. Each week included two 75-minute sessions focused primarily on listening and speaking activities. The participants were first-year students placed in the lowest groups based on their scores from a mock TOEIC test administered by the university.

Students registered for a section of the course based on their preferred time slots, resulting in the formation of a control group (N = 25) and an experimental group (N = 26). Both groups followed the same curriculum using *Interchange Intro, 5th Edition* (Cambridge University Press, 2021), covering thematic units such as self-introduction, describing places and moods, fashion, transportation, food, location, and telephone conversations. Each session focused on one unit, and the lesson structure involved pre-listening brainstorming and speaking, while post-listening activities focused on expressing opinions related to the topic.

In Week 2, students in both groups underwent a pre-test which included the TOEIC Speaking subtests "Respond to a Question" and "Express an Opinion," as well as the TOEIC Listening test. The same tests were administered again in Week 13 as post-tests. These two tasks were chosen from the TOEIC test format due to their frequent incorporation into classroom activities throughout the course. Additionally, prior to the intervention, only the experimental group completed a ChatGPT Literacy survey to assess their familiarity and baseline competence with the tool.

From Week 2 to Week 12, both groups received identical instructional content throughout the course, and all sessions were facilitated by the same instructor to maintain instructional consistency. To ensure uniformity, weekly speaking sessions were carefully controlled: both the duration of speaking practice and the tasks assigned were kept consistent across groups. That is, the weekly listening activities for both groups were based on the corresponding units from the textbook. The only variable that differentiated the two groups was the method of speaking practice. The control group engaged in traditional face-to-face peer conversations, whereas the experimental group replaced these with AI-mediated speaking practice using ChatGPT. Speaking activities were conducted as both pre-listening and post-listening tasks, with each session lasting approximately ten minutes (Synekop et al. 2024). Both groups discussed the same set of questions during their speaking sessions. This design enabled a meaningful comparison between traditional speaking formats and technology-enhanced instruction. During Week 5, for example, learners in the experimental group participated in speaking tasks guided by prompts designed to elicit personal reflection and future aspirations as the control group did in face-to-face interaction with

their peers. Example questions included: What did you want to be when you were younger? What do you want to be now? What skills or qualities are needed to get a job? Where would you like to work? What is your dream job, and why? Do you prefer working alone or with others? Would you like to work in a large, well-paying company? What is the most exciting aspect of your dream job? These questions encouraged students to practice expressing opinions, using descriptive language, and engaging in sustained dialogue, thereby simulating real-world communication scenarios through AI interaction.

After completing the course, the experimental group was also asked to answer post-ChatGPT literacy survey and reflect on their experience by describing the strengths, weaknesses, and suggestions for using ChatGPT as a learning tool. All assessments were used to evaluate the students' progress in speaking and listening skills, as well as to measure the educational effectiveness of ChatGPT-supported instruction.

3.3 Instruments

3.3.1 Pre- and post-English speaking & listening tests

To evaluate gains in English speaking and listening proficiency, this study utilized a standardized assessment based on the TOEIC. The listening test comprised a total of 30 items distributed across four parts: Part 1 (3 items), Part 2 (15 items), Part 3 (6 items), and Part 4 (6 items). The test duration was approximately 14 minutes, and identical pre- and post-tests were administered in Weeks 2 and 13, respectively, to facilitate direct performance comparisons.

Both tests were derived from the TOEIC speaking section featured in a commercially available TOEIC preparation book. The speaking test focused on two task types frequently practiced during regular class activities: *Respond to Questions* and *Express an Opinion*. The *Respond to Questions* task included the followings related to disposable items: (Q1) when do you usually use disposable items? (Q2) What kind of disposable items do you most often use? And (Q3) what is the advantage and disadvantage of using a disposable item? The *Express an Opinion* task required students to express their opinion: (Q4) some people prefer living in a smaller home in the city and others prefer a large home in the country. Which do you prefer and why?

The test required students to respond within specific time limits: 15 seconds for Q1 and 2, 30 seconds for Q3, and 60 seconds for Q4. Their spoken responses were recorded and submitted to the school's LMS system for assessment. The speaking assessments were administered during the second week (pre-test) and the thirteenth week (post-test) of the semester. The "Respond to questions" task is scored from 0 to 3 points per question. The "Giving an opinion" task is scored from 0 to 5 points.

To assess speaking test scores, two raters were involved: the researcher and a second evaluator, a university instructor with over 20 years of teaching experience. Speaking performance was evaluated using the TOEIC Speaking rubric, which includes criteria such as pronunciation, intonation and stress, grammar, vocabulary, cohesion, relevance of content, and completeness of content. To ensure consistency in scoring, the two raters first reviewed and discussed the assessment criteria together. They then evaluated approximately five sample students, compared their scores, and reached a consensus before proceeding with the full scoring process. After this norming session, each rater independently assessed the remaining students. The resulting reliability coefficient was 0.91, demonstrating a high level of internal agreement between the raters.

3.3.2 Pre- and post-questionnaires

To obtain comprehensive background information on the participants, a pre-experimental survey was administered. The questionnaire included items on participants' major, age, overseas experience, English CSAT (College Scholastic Ability Test) level, the areas of English they felt most confident in (up to two choices), the areas they most wished to improve (up to two choices), and the areas of English they would like to study using ChatGPT (up to two choices). In addition, participants were asked whether they had used ChatGPT before, and whether they had experience using it for learning purposes.

Following the initial survey, an additional ChatGPT Literacy Questionnaire was conducted for students in the experimental group. The initial focus of this study involved the ChatGPT literacy questionnaire, which comprised 34 items organized under five separate domains (Lee and Park 2024). The five factors are described as follows: The first factor, *Technical Proficiency (TP)* (#1-8), is factor emphasizes users' understanding of ChatGPT's underlying architecture, ability to manage technical aspects, and adaptability to system updates. It includes skills like troubleshooting, fine-tuning, integrating ChatGPT with other technologies, comparing with other models, and grasping NLP-related concepts. Users proficient in this area show confidence in not only using the tool but also explaining how it operates and evolves. The second factor, *Critical Evaluation (CE)* #9-16, centers on the ability to assess ChatGPT's responses critically. It covers evaluating accuracy, truthfulness, reliability, completeness, and bias. Learners are also expected to compare model versions and cross-reference AI outputs with other sources. Mastery in this factor reflects strong analytical judgment and skepticism when interpreting AI-generated content. The third factor, *Communication Proficiency (CP)* #17-23, reflects users' skill in interacting effectively with ChatGPT through precise, situation-aware language. It includes asking clear, vocabulary-rich questions, integrating technical terms, and tailoring interactions for decision-making or collaboration. Users strong in CP use ChatGPT not just as a Q & A tool, but as a conversational partner that supports strategic thinking. The fourth factor, *Creative Application (CA)* #24-29 focuses on the innovative use of ChatGPT for tasks like idea generation, storytelling, trend analysis, and productivity enhancement. It also includes using the tool to meet personal learning goals. Proficiency suggests that users view ChatGPT as a creative engine, helping them push boundaries in both personal and professional contexts. The last factor, *Ethical Competence (EC)* #30-34, involves awareness of privacy, legal concerns, responsible use, and the societal impact of AI interactions. Learners with high competence in this domain demonstrate a thoughtful approach to technology, considering not just what it can do, but how it should be used (see Appendix).

The purpose of the ChatGPT literacy questionnaire was to examine changes in the experimental group's literacy competencies related to ChatGPT before and after the instructional intervention. This survey was administered exclusively to students in the experimental group. Each item in the questionnaire was assessed using a six-point Likert scale, ranging from 1 ("Strongly disagree") to 6 ("Strongly agree").

In addition to the structured questions, participants were invited to provide open-ended responses reflecting their experiences using ChatGPT for English language learning. Specifically, they were asked to freely describe the strengths and weaknesses of ChatGPT-supported learning and to offer any suggestions for its improvement based on their personal experiences throughout the semester. Lastly, the post-survey included several items designed to capture students' perceptions of their learning experience. Participants were asked: (1) their preferred format for future English classes; (2) whether they intended to continue using ChatGPT to enhance their foreign language communication skills; and (3) whether they found ChatGPT-supported speaking and listening practice both effective and beneficial.

3.4 Data Analysis

This study aimed to examine the effectiveness of English instruction using ChatGPT by employing a range of assessment tools and analyzing the results. The assessment instruments included TOEIC listening and speaking tests, a ChatGPT Literacy questionnaire, and open-ended opinion surveys.

To evaluate changes in individual performance within each group (experimental and control), paired samples *t*-tests were conducted on pre- and post-test scores for both the TOEIC listening and speaking tests. Prior to conducting the *t*-test, a normality check was performed using the Shapiro-Wilk test, which is suitable for small to moderate sample sizes. The results showed that the data matched the assumption of normality ($p > .05$), which validated the use of parametric testing. Furthermore, to determine whether there were any differences in post-test scores between the groups, an independent samples *t*-test was first applied to pre-test scores to assess the baseline English proficiency. Since no significant differences were found at the pre-test level, a subsequent independent samples *t*-test was performed on the post-test scores to compare group outcomes. All quantitative data were analyzed using descriptive and inferential statistics with SPSS version 23.0.

Additionally, the experimental group completed a ChatGPT Literacy survey both before and after the intervention. This survey consisted of multiple-choice items rated on a 6-point Likert scale. Changes in literacy competence within the group were analyzed using paired samples *t*-tests.

Lastly, open-ended responses on the perceived strengths, weaknesses, and suggestions regarding the use of ChatGPT were collected. A thematic analysis was conducted, and the results were reported in terms of frequencies and percentages to better reflect learners' experiences. Frequencies of the remaining survey items were also aggregated and presented.

4. Results and Discussion

4.1 Effects of Generative AI on English Learning

4.1.1 Development of speaking and listening skills within groups

This study aimed to explore the impact of using ChatGPT as a supplementary tool in English speaking and listening activities. The primary objective was to investigate how ChatGPT-supported learning influences students' speaking and listening skills in English. To achieve this, participants were divided into two groups: a control group that received traditional instruction, and an experimental group that engaged in interactive speaking activities using ChatGPT. The study then analyzed changes in students' academic performance and learning attitudes to evaluate the effects of ChatGPT on English language acquisition.

Table 3. Result of Paired Sample *t*-tests: Control Group

Skill	Test	<i>M</i>	<i>SD</i>	<i>df</i>	<i>t</i>	<i>p</i>
Respond to questions	Pre	1.880	.634	24	-5.237	.000**
	Post	2.280	.597			
Express an opinion	Pre	2.280	.947	24	-6.532	.000**
	Post	2.920	1.038			
Listening	Pre	10.760	4.630	24	-6.422	.000**
	Post	14.280	4.098			

Note: ** $p < .01$

A series of paired sample *t*-tests were conducted to examine the performance differences between the pre-test and post-test scores within the control group ($N = 25$) across three language skill areas: responding to questions, expressing opinions, and listening comprehension (Table 3).

In the “respond to a question” speaking task, participants’ mean scores increased from $M = 1.880$ ($SD = 0.634$) in the pre-test to $M = 2.280$ ($SD = 0.597$) in the post-test. This difference was statistically significant, $t = -5.237$, $p < .01$. For the “express an opinion” skill, the pre-test mean was $M = 2.280$ ($SD = 0.947$), which improved to $M = 2.920$ ($SD = 1.038$) in the post-test. The difference was also statistically significant, $t = -6.532$, $p < .01$. Regarding listening comprehension, the participants showed notable improvement from a pre-test mean of $M = 10.760$ ($SD = 4.630$) to a post-test mean of $M = 14.280$ ($SD = 4.098$). The result was statistically significant as well, $t = -6.422$, $p < .01$.

Overall, these findings indicate that even within the control group receiving traditional instruction, the students in the control group demonstrated statistically significant improvement in all measured areas after the semester.

Table 4. Result of Paired Sample *t*-tests: Experimental Group

Skill	Test	<i>M</i>	<i>SD</i>	<i>df</i>	<i>t</i>	<i>p</i>
Respond to questions	Pre	1.866	.609	25	-6.584	.000**
	Post	2.423	.578			
Express an opinion	Pre	2.365	.625	25	-6.496	.000**
	Post	3.135	.742			
Listening	Pre	9.769	3.525	25	-8.104	.000**
	Post	13.539	4.032			

Note: ** $p < .01$

To evaluate the effect of ChatGPT-assisted instruction on students’ English speaking and listening skills, paired sample *t*-tests were conducted for the experimental group ($N = 26$) across three skill categories: responding to questions, expressing opinions, and listening comprehension.

In the “respond to a question” speaking task, participants’ scores significantly improved from a pre-test mean of $M = 1.866$ ($SD = 0.609$) to a post-test mean of $M = 2.423$ ($SD = 0.578$), $t = -6.584$, $p < .01$. For the “express an opinion” speaking task, the mean score increased from $M = 2.365$ ($SD = 0.625$) in the pre-test to $M = 3.135$ ($SD = 0.742$) in the post-test. This difference was statistically significant, $t = -6.496$, $p < .01$. In the listening comprehension test, the experimental group showed a notable rise in performance, with scores rising from $M = 9.769$ ($SD = 3.525$) to $M = 13.539$ ($SD = 4.032$). The improvement was statistically significant, $t = -8.104$, $p < .01$.

These results suggest that the integration of ChatGPT into speaking and listening activities led to significant gains in learners’ English proficiency across all assessed domains.

4.1.2 Comparative changes in speaking and listening skills between groups

The purpose of this study was to explore the impact of using the AI-supported language model, ChatGPT, on learners’ English listening and speaking skills. The primary research objective was to determine whether the integration of ChatGPT in classroom activities led to any significant differences in students’ language performance—specifically in the domains of listening and speaking—compared to a group that did not use ChatGPT.

To investigate this, an independent *t*-test was conducted to compare the post-test scores of the two groups in both listening and speaking. This statistical analysis allowed the researchers to examine whether the use of

ChatGPT had a significant effect on learners' oral communication competency.

Additionally, in order to establish the homogeneity of the two groups prior to the experiment, an independent *t*-test was performed on the pre-test scores. The analysis confirmed that there were no statistically significant differences between the groups in terms of their initial listening and speaking abilities. This validation ensured the reliability of the post-intervention comparison.

Table 5. Group Differences in the Pre-test

Skill	Group	<i>M</i>	<i>SD</i>	<i>df</i>	<i>t</i>	<i>p</i>
Respond to questions	Cont. (N=25)	1.880	.634	49	.084	.933
	Exp. (N=26)	1.866	.609			
Express an opinion	Cont. (N=25)	2.280	.947	49	-.381	.705
	Exp. (N=26)	2.365	.625			
Listening	Cont. (N=25)	10.760	4.630	49	.862	.393
	Exp. (N=26)	9.769	3.525			

Note: Cont. Control Group; Exp. Experimental Group

Prior to the intervention, an independent *t*-test was conducted to examine the homogeneity between the control and experimental groups across three language skill domains: responding to a question, expressing an opinion, and listening comprehension. The purpose of this analysis was to confirm whether any statistically significant differences existed between the groups.

As shown in Table 5, the results indicated no significant differences across all measured areas. For the skill "Respond to a question," the control group ($M = 1.880$, $SD = .634$) and experimental group ($M = 1.866$, $SD = .609$) showed negligible variation, $t = .084$, $p = .933$. Similarly, for "Express an opinion," the control group ($M = 2.280$, $SD = .947$) and the experimental group ($M = 2.365$, $SD = .625$) yielded a non-significant result, $t = -.381$, $p = .705$. In listening, the control group ($M = 10.760$, $SD = 4.630$) and the experimental group ($M = 9.769$, $SD = 3.525$) also showed no significant difference, $t = .862$, $p = .393$.

These results suggest that both groups were statistically equivalent at the outset of the study, thereby supporting the validity of the subsequent post-test comparisons.

Table 6. Group Differences in the Post-test

Skill	Group	<i>M</i>	<i>SD</i>	<i>df</i>	<i>t</i>	<i>p</i>
Respond to questions	Cont. (N=25)	2.280	.597	49	-.870	.388
	Exp. (N=26)	2.423	.578			
Express an opinion	Cont. (N=25)	2.920	1.038	49	-.852	.398
	Exp. (N=26)	3.135	.742			
Listening	Cont. (N=25)	14.280	4.098	49	.651	.518
	Exp. (N=26)	13.539	4.032			

Note: Cont. Control Group; Exp. Experimental Group

To determine the impact of ChatGPT use on learners' speaking and listening performance, an independent samples *t*-test was conducted to compare post-test scores between the control and experimental groups across three skill areas: responding to a question, expressing an opinion, and listening comprehension.

As shown in Table 6, no statistically significant differences were observed between the two groups across all measured domains. For "Respond to a question," the experimental group ($M = 2.423$, $SD = .578$) scored slightly higher than the control group ($M = 2.280$, $SD = .597$), yet the difference was not significant, $t = -.870$, $p = .388$. Similarly, in "Express an opinion," the experimental group ($M = 3.135$, $SD = .742$) outperformed the control group

($M = 2.920$, $SD = 1.038$), but this difference was also not statistically significant, $t = -.852$, $p = .398$. In the “Listening” section, the control group ($M = 14.280$, $SD = 4.098$) scored slightly higher than the experimental group ($M = 13.539$, $SD = 4.032$), yet the difference remained non-significant, $t = .651$, $p = .518$.

The findings revealed that although the experimental group showed slightly better performance in speaking tasks, ChatGPT use did not result in statistically significant improvements in speaking or listening skills compared to the control group. This contrasts with earlier studies by Cheon (2023) and Noh (2024). Cheon (2023) reported notable gains in vocabulary acquisition and cultural appropriateness among ChatGPT users, while Noh found significant improvements in vocabulary, pronunciation, and fluency, along with reduced psychological barriers to speaking.

According to Huang (2024) and Yang (2025), interaction with ChatGPT promotes dynamic practice and personalized feedback, thereby enhancing speaking proficiency through well-structured prompts that elicit meaningful responses. In light of this, it is plausible that the present study might have yielded more significant results had learners been provided with well-structured prompts designed to facilitate deeper engagement. This suggests the need for educators to develop and implement specific strategies for crafting prompts that guide learners toward meaningful interaction when using AI tools such as ChatGPT.

4.2 Changes in ChatGPT Literacy

To examine how students perceived AI-supported speaking activities as a substitute for face-to-face conversations during class, pre- and post-surveys were conducted. The survey consisted of 34 items designed to capture a range of learner perspectives, organized into five sub-factors.

Table 7 presents the results of the analysis, highlighting any significant differences in mean scores across the sub-factors between the pre- and post-survey stages. Specifically, the experimental group responded to each item before and after using the program, and the mean scores along with standard deviations were measured. These results provide a systematic analysis of how students’ perceptions shifted after engaging in AI-supported speaking activities.

Table 7. Results of Pre- & Post-Survey: ChatGPT Literacy

Factor	Survey	<i>M</i>	<i>SD</i>	<i>df</i>	<i>t</i>	<i>p</i>
Factor 1	Pre	2.776	.795	23	-3.711	.001**
	Post	3.563	.572	23		
Factor 2	Pre	2.839	.757	23	-3.849	.001**
	Post	3.432	.584	23		
Factor 3	Pre	2.893	.766	23	-4.142	.000**
	Post	3.548	.498	23		
Factor 4	Pre	3.153	.782	23	-3.648	.001**
	Post	3.729	.540	23		
Factor 5	Pre	3.333	.974	23	-1.983	.059
	Post	3.650	.630			

Note: Factor 1: Technical Proficiency (TP), Factor 2: Critical Evaluation (CE), Factor 3: Communication Proficiency (CP), Factor 4: Creative Application (CA), Factor 5: Ethical Competence (EC)

** $p < .01$

The first factor, *Technical Proficiency (TP)*, demonstrated a significant improvement in learners’ understanding of ChatGPT’s operational mechanisms. The pre-intervention mean score was 2.776 ($SD = 0.795$), which increased to 3.563 ($SD = 0.572$) post-intervention. A paired *t*-test indicated a statistically significant difference ($t = -3.711$,

$p = .001$), suggesting that learners enhanced their ability to navigate ChatGPT's interface, apply prompt engineering, and utilize its core functionalities effectively.

The second factor, *Critical Evaluation (CE)*, also showed notable development, with mean scores rising from 2.839 ($SD = 0.757$) to 3.432 ($SD = 0.584$). This shift was statistically significant ($t = -3.849$, $p = .001$), indicating that learners became more adept at assessing the accuracy, reliability, and potential bias in ChatGPT-generated outputs. Furthermore, their ability to cross-reference information and apply corrective judgment improved substantially.

In regard to *Communication Proficiency (CP)*, learners exhibited the most pronounced growth among all dimensions, with pre-post means increasing from 2.893 ($SD = 0.766$) to 3.548 ($SD = 0.498$). The t -value of -4.142 and p -value of .000 point to a highly significant improvement. This enhancement reflects learners' refined questioning strategies, fluency in dialogue management, and sensitivity to contextually appropriate linguistic choices during AI interaction.

The fourth factor, *Creative Application (CA)*, was also positively impacted, with the pre-intervention mean of 3.153 ($SD = 0.782$) rising to 3.729 ($SD = 0.540$) after the course. The statistical outcome ($t = -3.648$, $p = .001$) confirms meaningful gains in learners' capacity to harness ChatGPT for ideation, storytelling, and content generation. These results suggest a shift from passive tool use to active creative collaboration.

Lastly, *Ethical Competence (EC)* showed a modest increase, from 3.333 ($SD = 0.974$) to 3.650 ($SD = 0.059$). However, this change did not reach statistical significance ($t = -1.983$, $p = .059$), implying that while learners' awareness of ethical considerations—such as data privacy and intellectual property—slightly improved, there was insufficient behavioral or practical uptake. This finding underscores the need for more structured and experiential ethics instruction within future iterations of the course.

Overall, the results affirm that instruction centered on ChatGPT can lead to meaningful improvements in technical, evaluative, communicative, and creative literacy, with particularly strong effects observed in communication and critical thinking domains. These findings appear to align with the results reported by Kim et al. (2025). Specifically, their study demonstrated that integrating ChatGPT into EFL classrooms over the course of a semester led to notable improvements in students' ChatGPT literacy. This, in turn, not only contributed to the enhancement of their English language skills but also supported the development of competencies relevant to real-life communication and problem-solving (Lee and Park 2023).

4.3 Learners' Views on the Integration of ChatGPT into English Learning

To gain deeper insights into the experimental group's perceptions of AI-supported instruction, open-ended survey items were administered. These open-ended items—addressing perceived benefits, drawbacks, and future suggestions—were formulated to prompt learners to reflect on their semester-long experiences of using ChatGPT in spoken English practice.

Upon collection of the responses, a thematic analysis was conducted whereby individual statements were classified according to recurring themes. Frequencies and percentages were then calculated for each thematic category to identify predominant patterns and the relative emphasis placed by participants across dimensions. Of the 26 students in total, two who failed to complete either the pre- or post-survey were excluded from the analysis.

Table 8 presents the reported advantages of using ChatGPT as perceived by the learners throughout the semester. The most prominent benefit, cited by 19.4% of participants, was the improvement in speaking skill. Learners emphasized that they were able to speak more naturally and fluently.

Table 8. Advantages of Using ChatGPT

	Category	Example Description	#	%
1	Improved speaking skill	Speaking more fluently/ speaking more naturally	6	19.4
2	Better communicative skills	Responding to ChatGPT / integrated speaking & listening	4	12.9
3	Psychological comfort & reduced pressure	No shame in mistakes / easier to talk than with people/ gained confidence	4	12.9
4	Detailed & diverse responses	Multiple perspectives / clear explanations	3	9.7
5	No time or space limitations	Fast communication / learning alone / accessible anytime	3	9.7
6	Practical speaking practice	Continuous practice / follow-up question flow	3	9.7
7	Cultural & informational exploration	Discovering foreign culture / researching topics of interest	2	6.5
8	Use of voice features	Useful listening / feel like real conversation	2	6.5
9	Improved perception of ChatGPT	Realizing its capabilities / reduced aversion	2	6.5
10	Unexpected fun element	Surprising or amusing answers	1	3.2
11	Conversation logs & feedback possible	Able to save conversations and get feedback	1	3.2

Equally noted (12.9%) were gains in listening and interaction skills, attributed to engaging in real-time responses with the AI, which fostered integrated development in both speaking and listening. Another 12.9% highlighted psychological comfort, reduced communicative pressure, and gained confidence; learners expressed that speaking with ChatGPT felt easier than human interactions, primarily due to the absence of judgment and fear of embarrassment.

Additional advantages included receiving detailed and diverse responses (9.7%), which enabled exposure to multiple perspectives and facilitated comprehension through clear explanations and concise summaries. Similarly, learners appreciated the absence of temporal or spatial constraints (9.7%), noting the convenience of solo learning and the accessibility of practice at any time.

Practical aspects of speaking practice were also valued (9.7%), with learners mentioning the seamless flow of follow-up questions and the opportunity for continuous engagement. Cultural and informational exploration (6.5%) was another benefit, as ChatGPT introduced learners to foreign cultures and provided insights into various topics of interest.

The voice feature contributed to improved listening experiences and simulated authentic conversation (6.5%), while an improved perception of GPT itself (6.5%) was noted by those who developed greater trust in the technology and overcame initial aversion. Less frequently mentioned, but still notable, were unexpected elements of enjoyment (3.2%) and the ability to review conversation logs for feedback purposes (3.2%).

As a result, the positive responses from students regarding their use of ChatGPT align with findings from previous studies. Specifically, Liu and Ma (2024) highlighted that learners' perceptions—such as perceived usefulness and satisfaction—play a critical role in shaping their attitudes and intentions toward ChatGPT use (Choi and Choe 2024, Kim and Kim 2024). Similarly, Synekop et al. (2024) reported that learners acknowledged ChatGPT's potential to enhance communicative competence and emphasized the importance of integrating AI tools into English education to support language acquisition and respond to evolving educational demands. These findings suggest that ChatGPT contributes not only to learners' psychological engagement but also to their linguistic development (Huang 2024).

Table 9 summarizes the perceived disadvantages of using ChatGPT for spoken English practice, based on learners' semester-long reflections. The most commonly cited issue (26.7%) involved speech recognition errors, particularly the AI's failure to accurately detect pronunciation, interference from ambient noise, and

misinterpretation of questions.

Table 9. Disadvantages of Using ChatGPT

	Category	Example Description	#	%
1	Speech recognition errors	Failed pronunciation detection / ambient noise / misheard questions	8	26.7
2	Incorrect info & disrupted flow	Wrong answers / broken conversational flow / error responses	5	16.7
3	Technical glitches & operational issues	App crashes / malfunction / restart required	3	10.0
4	Limited question clarity or comprehension	Misunderstood the questions/ vocabulary	2	6.7
5	Response length or complexity issues	Answer too long / explained in a difficult way / user requests not met	2	6.7
6	Lack of casual conversation	Missing small talk / too much information	2	6.7
7	Poor mixed-language recognition (Korean + English)	Topic confusion when switching languages	2	6.7
8	Repetitiveness or creativity gaps	Similar answers repeated / lack of spontaneous conversation	2	6.7
9	Limited grammar feedback or accuracy check	Unclear if user's sentences were correct / lacking feedback tools	2	6.7
10	Unfamiliarity with usage	Unclear how to use / felt unfamiliar or difficult to engage	2	6.7
11	AI cannot express opinions	No opinion given / lack subjective perspective	1	3.3

Another significant concern (16.7%) was the disruption caused by incorrect responses and broken conversational flow. Learners reported that inaccurate information or technical error responses undermined the continuity and effectiveness of communication. Technical glitches (10.0%)—including application crashes, malfunctions, and forced restarts—also detracted from the overall experience.

A range of limitations were noted by 6.7% of participants each. These included difficulties in expressing questions clearly due to limited vocabulary and AI's inability to comprehend the user's intent; issues with lengthy or overly complex answers that failed to meet learner expectations; and a perceived lack of casual, small-talk-style conversation, with some responses feeling too rigid or information-heavy.

Additionally, learners pointed out problems with mixed-language recognition (Korean-English), resulting in topic confusion when switching languages. Repetitive output and limited creativity were highlighted, along with the lack of spontaneous conversational variety. Learners also expressed dissatisfaction with grammar feedback mechanisms, noting the absence of tools to check sentence accuracy or provide corrective input.

Further disadvantages included a sense of unfamiliarity or difficulty in engaging with the AI (6.7%), particularly among those less accustomed to digital tools. Finally, one learner (3.3%) remarked on the AI's inability to express subjective opinions, which was perceived as a limitation in dynamic or opinion-based discussions.

Despite ongoing advancements in AI technology, persistent issues such as speech recognition errors and technical glitches continue to surface. These limitations not only disrupt the flow of interaction but also hinder the overall reliability and usability of AI tools in educational settings, reinforcing the need for strategic support and user awareness. The findings of this study align with the claims of Kim et al. (2025), suggesting that learners must employ strategies such as follow-up questioning, refined prompt usage, and cross-verification with reliable sources to ensure the trustworthiness of AI-generated responses. These results highlight the need for instructional support that fosters learners' information literacy and critical thinking skills in AI-integrated educational contexts. Furthermore, consistent with the study by Oh (2023), concerns about the quality and reliability of information

underscore the importance of pedagogical guidance, including source comparison, critical evaluation, and clear ethical frameworks. This implies that when utilizing AI tools like ChatGPT, learners should be equipped not merely to accept information passively, but to engage with it critically and make informed judgments.

Table 10. Suggestion for Using ChatGPT

	Category	Example Description	#	%
1	Satisfied with current usage & prefer to keep it	Like current format / no complaints / want to continue as is	7	26.9
2	Need to explore when ChatGPT is most effective	Unsure when ChatGPT is useful / need guidance	4	15.4
3	Improve questioning strategy	Use more specific instructions / design better questions	4	15.4
4	Utilize ChatGPT to explore culture or build writing	Access cultural info / use as quiz tool / help with composition	3	11.5
5	Improve voice recognition environment	Use headphones / reduce surrounding noise / try at home	2	7.7
6	Feedback and review functions desired	Want ChatGPT to give feedback / build follow-up questioning skills	2	7.7
7	Use GPT for general casual conversation	Practice small talk / everyday conversational skills	2	7.7
8	Provide extra time or task for shy learners	Assign ChatGPT practice as homework / more time for quiet students	1	3.8
9	Improve recognition accuracy	Accuracy when listening should be better	1	3.8

Table 10 outlines learner suggestions for optimizing the use of ChatGPT in spoken English practice. A majority of learners (26.9%) expressed satisfaction with the current use of ChatGPT, indicating that the existing format met their needs and that they preferred to continue without major changes. Others (15.4%) recommended further exploration into when and how ChatGPT is most effective, highlighting a need for clearer guidelines and structured usage scenarios. An equal proportion (15.4%) emphasized improving the questioning strategy—encouraging users and instructors to craft more specific prompts and well-designed questions to maximize learning outcomes.

Several participants (11.5%) saw potential in leveraging ChatGPT for cultural exploration and writing development, including applications such as accessing cultural information, using quizzes, and receiving composition support. Enhancing the voice recognition environment (7.7%) was another area of interest, with suggestions to use noise-cancelling headphones or practice in quieter locations to improve interaction quality.

Learners also requested feedback and review capabilities (7.7%), noting that GPT could support follow-up question skills or help evaluate spoken output. Practicing casual conversation (7.7%) was mentioned as a desirable feature, especially for building real-world communication fluency in informal contexts.

Less frequently noted but still valuable, one learner (3.8%) suggested allocating extra time or assigning GPT-based tasks to accommodate shy learners who might benefit from extended practice periods. Lastly, improving recognition accuracy (3.8%)—especially in terms of listening comprehension—was suggested as a technical enhancement.

In summary, students proposed a range of suggestions to enhance the effective use of ChatGPT as a learning tool. While most expressed satisfaction with the current usage and wished to maintain it, they also emphasized the need to develop questioning strategies to maximize its effectiveness. Some students suggested using ChatGPT to explore cultural topics and improve writing skills, highlighting its potential to support broader learning goals. Additionally, there were recommendations to improve the voice recognition environment and accuracy to enable smoother interactions, as well as to incorporate feedback and review functions to enhance learning outcomes. For more introverted learners, assigning tasks and providing structured opportunities for engagement were seen as

ways to encourage self-directed learning. Taken together, these suggestions point toward a more effective and comprehensive use of ChatGPT in educational contexts.

In response to the survey item asking students about their preferred format for future English classes, 13 students expressed a preference for a blended approach combining ChatGPT and traditional instruction, while 10 students preferred utilizing ChatGPT alone. Only one student indicated a preference for the traditional format exclusively. When asked whether they intended to continue using ChatGPT to improve their foreign language communication skills, 22 students responded affirmatively, whereas two stated they would not use it extensively. Overall, the majority of students perceived their experience with ChatGPT-supported speaking and listening practice as both effective ($N = 22$) and beneficial ($N = 23$), highlighting its pedagogical value in EFL contexts.

As a result, after experiencing AI-supported English learning, most students showed a clear preference for this approach over traditional methods and expressed a strong intention to continue using ChatGPT to improve their communication skills. This suggests that students are not only learning how to effectively integrate AI into their language learning process, but also developing self-directed learning abilities (Lee and Lee 2021). Given that self-directed learning constitutes a multifaceted trait involving goal-setting, autonomous decision-making, and reflective evaluation (Garrison 1997), the study's findings—such as learners actively engaging with ChatGPT and seeking feedback independently can be interpreted as indicators of enhanced self-regulation and learner autonomy. These behaviors reflect important aspects of self-directed learning and indicate that AI-integrated environments may promote its growth.

5. Conclusion

This study aimed to investigate whether Korean university students—who have limited opportunities for direct interaction with native English speakers in an English as a Foreign Language (EFL) environment—could improve their English communication skills through the use of an AI-powered tool, ChatGPT. Furthermore, by incorporating ChatGPT continuously over the course of a semester, the study sought to explore the extent to which learners were able to develop the ChatGPT literacy required for effective use of ChatGPT in the AI era, as well as to examine their perceptions and attitudes formed through this experience.

To address the first research question—To what extent does AI-supported speaking practice enhance learners' English speaking and listening skills in an EFL context—this study examined the impact of ChatGPT-supported learning by comparing an experimental group utilizing ChatGPT with a control group receiving traditional instruction. Within-group analyses revealed that both groups significantly improved in responding to questions, expressing an opinion, and listening comprehension over the semester. However, between-group comparisons showed no statistically significant differences in post-test scores across these skills. Although the experimental group using ChatGPT demonstrated slightly higher gains in speaking, the results suggest that ChatGPT did not produce superior outcomes relative to traditional methods. Nonetheless, the overall improvement in both groups indicates that structured speaking and listening activities—whether AI-supported or not—can enhance language proficiency.

Contrary to expectations, the study found no statistically significant improvement in speaking or listening skills from using ChatGPT, despite slightly better speaking performance in the experimental group. According to Lee (2024), learners' proficiency levels influence both the frequency of prompt usage and the degree of sustained interaction with ChatGPT. In other words, higher-level learners tend to use prompts more actively and engage in more dynamic exchanges, whereas lower-level learners show limited use and interaction. Given that the

participants in the present study were relatively low-level English learners, it is plausible that their limited use of prompt strategies and lack of sustained interaction contributed to the absence of significant learning outcomes. This may help explain why the results did not meet initial expectations. Researchers (Huang 2024, Yang 2025) emphasized that well-structured prompts and interactive feedback with ChatGPT can enhance language proficiency. However, the present study did not sufficiently provide learners with guidance or strategies on how to effectively use prompts, which may have limited the potential impact of ChatGPT. This highlights the need for educators to design and implement clear instructional strategies for prompt use, enabling learners to engage in more meaningful interactions with AI tools like ChatGPT.

Regarding the second research question—How does engaging in AI-supported language learning contribute to the development of ChatGPT literacy among university students—learners demonstrated meaningful progress across five core competencies through their experience with ChatGPT-integrated instruction. In terms of technical proficiency, they became more adept at navigating the interface, crafting effective prompts, and utilizing core features. Their ability to critically evaluate AI-generated content also improved, particularly in assessing accuracy, reliability, and bias. Communication proficiency saw the most remarkable development, as learners refined questioning strategies, enhanced dialogue management, and responded more appropriately to contextual cues during interactions. Creative application skills were strengthened as learners began using ChatGPT more proactively for idea generation and storytelling. Lastly, although ethical competence showed a slight improvement, it did not translate into significant behavioral change, underscoring the importance of incorporating more structured ethics education in future curricula.

In summary, the findings suggest that incorporating ChatGPT into instruction can significantly enhance various aspects of ChatGPT literacy. These outcomes are consistent with those reported by Kim et al. (2025), whose study showed that semester-long integration of ChatGPT in EFL classrooms led to marked improvement in students' ChatGPT literacy. This improvement not only can boost their English proficiency but also foster skills essential for authentic communication and effective problem-solving (Lee and Park 2023)

Lastly, regarding learners' perceptions and experiences of using ChatGPT for speaking practice throughout the semester, students reported a broad range of benefits. These included enhanced speaking skill, improved listening and interaction skills, psychological comfort and confidence, and more accessible learning opportunities. Students' favorable responses to ChatGPT align with previous research showing that learners' perceptions significantly shape their attitudes and intentions to engage with the tool (Choi and Choe 2024, Kim and Kim 2024, Liu and Ma 2024). Likewise, students recognized ChatGPT's value in improving communication skills and stressed the need to include AI tools in English education to support language learning and adapt to changing educational needs (Huang 2024, Synekop et al. 2024).

While AI tools offer notable benefits, several limitations in using AI tools include both technical and language challenges: Speech recognition errors and system glitches often lead to inaccurate information and disrupt the flow of interaction. AI may struggle to comprehend unclear questions, produce overly long or complex responses, and maintain natural, casual conversation. These challenges continue to undermine the reliability and effectiveness of AI in educational contexts. In response, this study echoes Kim et al. (2025), who stress the importance of user strategies like precise prompt formulation and cross-checking information with credible sources. It also aligns with Oh (2023), underscoring the need for pedagogical support that cultivates critical thinking, ethical reasoning, and informed engagement with AI-generated content.

Based on their experience, participants offered several recommendations, such as providing clearer usage instructions, improving question formulation techniques, namely prompt literacy (Hwang 2023, Yang 2025), enhancing voice interaction quality, and creating more opportunities for informal, conversational exchanges.

Survey responses further revealed a strong preference for blended instruction formats and continued use of ChatGPT, with most students perceiving the tool as both effective and pedagogically valuable. In sum, after using AI-supported English learning, most students preferred it over traditional methods and planned to keep using ChatGPT to boost communication skills. This reflects growing self-directed learning and improved AI literacy (Lee and Lee 2021).

The findings of this study underscore the pedagogical value of integrating AI-supported language learning into foreign language education. Beyond enhancing learners' emotional stability and linguistic proficiency, such integration cultivates AI literacy—an essential competency in the digital era. Participants who experienced AI-assisted instruction reported a heightened willingness to engage voluntarily with the technology and expressed a sustained interest in its continued application in English language courses. These outcomes indicate that AI can play a pivotal role in motivating learners, sustaining their engagement, and fostering deeper learning immersion. Therefore, rather than serving merely as a supplementary tool, AI should be positioned as an educational medium that advances learner autonomy, problem-solving ability, and self-directed learning. Consequently, the effective implementation of AI in instructional contexts demands more than the simple adoption of technology; it requires carefully designed pedagogical frameworks and learner-centered approaches aligned with educational objectives. By taking these considerations into account, AI-supported language learning can function as an agent of change in the future of foreign language education.

This study has several limitations that should be acknowledged. First, the research was conducted with students from one single university, whose overall English proficiency was relatively low, thereby constraining the generalizability of the findings to broader populations. Although ChatGPT was utilized as a conversational tool in the instructional intervention, the range of tasks integrated into the program remained narrow, limiting the exploration of its full pedagogical potential. Additionally, speaking performance was evaluated using only holistic scoring methods, without conducting a more subtle analysis based on analytic rubrics, which could have provided deeper insights into specific communicative competencies.

To address these limitations, future research should consider involving learners across various proficiency levels and institutional contexts to better represent the diversity of EFL learners. Moreover, qualitative investigations into how learners engaged with AI literacy in practice would enrich the understanding of cognitive and behavioral dimensions of AI-supported learning. To better understand the impact of AI-assisted learning, it is essential to first teach students prompt literacy and effective questioning strategies, then examine how these skills influence their communicative competence. Additionally, engaging students in a variety of AI-supported tasks would allow researchers to observe changes in their self-directed learning abilities, offering deeper insight into the educational potential of AI tools. Such approaches could yield more meaningful educational implications and contribute to the development of effective instructional methods in AI-integrated language education.

References

- Baidoo-Anu, D. and L. Owusu Ansah. 2023. Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. *Journal of AI* 7(1), 52-62.
- Bao, W. and R. Li. 2023. ChatGPT and AI: The game changer for education. *Journal of Information Technology Case and Application Research* 25(3), 277-304.

- Baskara, R. and M. Mukarto. 2023. Exploring the implications of ChatGPT for language learning in higher education. *Indonesian Journal of English Language Teaching and Applied Linguistics* 7(2), 343-358.
- Bok, E. and Y. Cho. 2023. Examining Korean EFL College students' experiences and perceptions of using ChatGPT as a writing revision tool, *Journal of English Teaching through Movies and Media* 24(4), 5-27.
- Cha, S. 2023. Exploring the utilization of ChatGPT in elementary English education. *Multimedia-Assisted Language Learning* 26(3), 130-150.
- Cheon, S.-M. 2023. Educational application of ChatGPT: Its impact on Korean university students' English speaking skills. *The Journal of Linguistic Science* 107, 469-496.
- Choe, Y. 2023. Exploring ChatGPT's impact on the English summary writing of pre-service English teachers. *Multimedia-Assisted Language Learning* 26(2), 104-132.
- Choi, S. and Y. Choe. 2024. Exploring the impacts of AI-based English instruction on Korean middle school EFL students' AI literacy and perceptions. *Multimedia-Assisted Language Learning* 27(1), 96-130.
- Cox, A. 2023. How artificial intelligence might change academic library work: Applying the competencies literature and the theory of the professions. *Journal of the Association for Information Science and Technology* 74(3), 367-380.
- Crosthwaite, P. and V. Baisa. 2023. Generative AI and the end of corpus-assisted data-driven learning? Not so fast! *Applied Corpus Linguistics* 3, Article 100066.
- Dignum, V. 2019. *Responsible Artificial Intelligence: How to Develop and Use AI in a Responsible Way*. Springer.
- Garrison, D. R. 1997. Self-directed learning: Toward a comprehensive model. *Adult Education Quarterly* 48(1), 18-33.
- Gass, S. M., J. Behney and L. Plonsky. 2020. *Second language acquisition: An introductory course* (5th ed.). Routledge.
- Gilster, P. 1997. *Digital Literacy*. Wiley Computer Publishing.
- Guo, K. and D. Wang. 2023. To Resist it or to Embrace it? Examining ChatGPT's Potential to Support Teacher Feedback in EFL Writing. *Education and Information Technologies*.
- Huang, J. 2024. Enhancing EFL speaking feedback with ChatGPT's voice prompts. *International Journal of TESOL Studies* 6(3), 4-23.
- Hwang, Y. 2023. The emergence of generative AI and PROMPT literacy: Focusing on the use of ChatGPT and DALL-E for English education. *Journal of the Korea English Education Society* 22(2), 263-288.
- Im, H.-J. 2023. A study on college students' perspectives and attitudes toward the use of ChatGPT in English classes. *Culture & Convergence* 45(9), 1335-1342.
- Jeon, J. and S. Lee. 2023. Large language models in education: A focus on the complementary relationship between human teachers and ChatGPT. *Educ Inf Technol* 28, 15873-15892.
- Jeong, N.-S. 2024. Exploring the effects of ChatGPT on university students' English writing skills and their perceptions. *Multimedia-Assisted Language Learning* 27(1), 78-95.
- Kim, H.-S. and N.-Y. Kim. 2024. Exploring the effects of ChatGPT on video-making projects in an EFL course. *Asia-pacific Journal of Convergent Research Interchange* 10(6), 737-750.
- Kim, H.-S., N.-Y. Kim, Y. Kim and D.-J. Son. 2024. Recent trends in using ChatGPT for English and Korean language education. *Multimedia-Assisted Language Learning* 27(4), 45-69.
- Kim, N.-Y., Y. Cha and H.-S. Kim. 2025. Exploring ChatGPT literacy in EFL reading education. *Multimedia-Assisted Language Learning* 28(1), 39-60.
- Kim, R. 2024. Effects of ChatGPT on Korean EFL learners' main-idea reading comprehension via top-down. *Language Research* 60(1), 83-106.

- Kohnke, L., B. L. Moorhouse and D. Zou. 2023. ChatGPT for language teaching and learning. *RELC*, 1-14.
- Lee, J. S. and K. Lee. 2021. The role of informal digital learning of English and L2 motivational self-system in foreign language enjoyment. *British Journal of Educational Technology* 52(1), 358-373.
- Lee, S. 2024. Exploring ChatGPT's potential as a supplementary feedback tool in English writing classes. *Foreign Languages Education* 31(2), 27-54.
- Lee, S. and G. Park. 2023. Exploring the impact of ChatGPT literacy on user satisfaction: The medicating role of user motivations. *Cyberpsychology, Behavior, and Social Networking* 26(12), 913-918.
- Lee, S. and G. Park. 2024. Development and validation of ChatGPT literacy scale. *Current Psychology* 43, 18992-19004.
- Liu, G. L., R. Darvin and C. Ma. 2024. Exploring AI-mediated informal digital learning of English (AI-IDLE): a mixed-method investigation of Chinese EFL learners' AI adoption and experiences. *Computer Assisted Language Learning* 1-29.
- Liu, M., L. J. Zhang and C. Biebricher. 2024. Investigating students' cognitive processes in generative AI-assisted digital multimodal composing and traditional writing. *Computers and Education* 211, 104977, 1-21.
- Liu, W. 2025. Language teacher AI literacy: insights from collaborations with ChatGPT. *Journal of China Computer-Assisted Language Learning*. 1-30.
- Lund, B., D. Agbaji and Z. A. Teel. 2023. Information literacy, data literacy, privacy literacy, and ChatGPT Technology literacies align with perspectives on emerging technology adoption within communities. *Human Technology* 19(2), 163-177.
- Ma, Q., P. Crosthwaite, D. Sun and D. Zou. 2024. Exploring ChatGPT literacy in language education: A global perspective and comprehensive approach. *Computers and Education: Artificial Intelligence* 7, Article 100278.
- McCallum, D. 2023. Generative AI and ChatGPT: Applications, challenges, and AI-human collaboration. *Journal of Information Technology Case and Application Research* 25(3), 277-304.
- Mizumoto, A. and M. Eguchi. 2023. Exploring the potential of using an AI language model for automated essay scoring. *Research Methods in Applied Linguistics* 2(2), 100050.
- Noh, Y. 2024. The impact of ChatGPT-assisted learning on English speaking proficiency of Korean university students. *The New Studies of English & Literature* 89, 89-114.
- Oh, S. 2023. A study on the case of using ChatGPT & learners' perceptions in college liberal arts writing. *Korean Journal of General Education* 17(3), 11-23.
- Raschke, C. A. 2003. *The Digital Revolution and the Coming of the Postmodern University*. Routledge.
- Sharples, M. 2023. Towards social generative AI for education: theory, practices and ethics. *Learning: Research and Practice* 9(2), 159-167.
- Shim, D. and H. Kim. 2024. The effects of using ChatGPT as a feedback tool on high school students' English writing. *Multimedia-Assisted Language Learning* 27(1), 25-50.
- Shin, H.-C. and J.-H. Lee. 2024. A comparative study of feedback from native English teachers and ChatGPT on Korean university students' English composition with student perceptions. *Multimedia-Assisted Language Learning* 27(3), 106-127.
- Synekop, O., I. Lytovchenko, Y. Lavrysh and V. Lukianenko. 2024. Use of ChatGPT in English for engineering classes: Are students' and teachers' views on its opportunities and challenges similar? *International Journal of Interactive Mobile Technologies (iJIM)* 18(3), 129-146.
- Yang, H. 2025. EFL student engagement with ChatGPT in college reading classes via prompts and perceptions. *Korean Journal of English Language and Linguistics* 25, 686-706.

Yim, I. H. Y. and J. Su. 2024. Artificial intelligence (AI) learning tools in K-12 education: A scoping review. *Journal of Computers in Education* 12, 93-131.

Young, J. C. and M. Shishido. 2023. Investigating Open AI's ChatGPT potentials in generating Chatbot's dialogue for English as a foreign language learning. *International Journal of Advanced Computer Science and Applications* 14(6), 65-72.

Examples in: English

Applicable Languages: English, L2

Applicable Level: Tertiary