NOVITAS-ROYAL
RESEARCH ON YOUTH AND LANGUAGE

2024, 18(2), 167–182
Novitas-ROYAL (Research on Youth and Language)
ISSN: 1307–4733

Research Article

ChatGPT Integration in EFL Education: A Path to Enhanced Speaking Self-Efficacy

Cennet YILDIZ[1]

[1] Uşak University, Uşak, TÜRKİYE
altinercennet@gmail.com
ORCID: 0000-0001-8930-4913

Article information

Submission	17/07/2024	Revision received	11/09/2024
Acceptance	13/09/2024	Publication date	22/10/2024

Keywords:
Speaking
Self-efficacy
ChatGPT
Mixed-method

Abstract: This study explored the effect of ChatGPT on the speaking self-efficacy level of English as foreign language learners. The research followed the mixed method approach and involved a total of 65 participants who were randomly assigned to the experimental and control groups. The experimental group used the ChatGPT for communicative activities, while the control group was engaged in a regular EFL classroom. The data collection procedure included the administration of the EFL speaking self-efficacy scale and an interview with the experimental group as optional participation. The t-test on independent samples showed non-statistically significant differences in speaking self-efficacy levels of the two groups before the intervention. Nevertheless, the covariance (ANCOVA) analysis after ten intervention sessions revealed noteworthy enhancements in speaking self-efficacy among learners in the experimental group. The interviews revealed that engaging with ChatGPT facilitated learners in developing confidence in their conversational skills, managing their stress levels, and deriving pleasure from the learning experience. The results indicated that interactions facilitated by AI chatbots, like ChatGPT, fostered a conducive learning environment for learners, leading to increased confidence in speaking.

Anahtar Sözcükler:
Konuşma
Öz yeterlik
ChatGPT
Karma yöntem

Yabancı Dil Olarak İngilizce Eğitiminde ChatGPT Entegrasyonu: Konuşma Öz Yeterliğini Artırma Yolu

Özet: Bu çalışmada, ChatGPT'nin İngilizceyi yabancı dil olarak öğrenenlerin konuşma öz-yeterlik düzeyi üzerindeki etkisi araştırıldı. Araştırmada karma yöntem yaklaşımı izlendi ve toplam 65 katılımcı deneysel ve kontrol gruplarına rastgele atanmıştır. Deneysel grup iletişimsel faaliyetler için ChatGPT aracını kullanırken, kontrol grubu sınıf içi aktivitelere devam etti. Veri toplama prosedürü, İngilizce Konuşma Öz-Yeterlik Ölçeği uygulamasını ve deneysel gruptaki katılımcılarla isteğe bağlı olarak yapılan bir mülakatı içerdi. Bağımsız örnekler üzerinde yapılan t-testi, müdahaleden önceki iki grup arasında konuşma öz-yeterlik düzeylerinde istatistiksel olarak anlamlı farklılıklar olmadığını gösterdi. Bununla birlikte, on müdahale oturumundan sonra gerçekleştirilen varyans analizi (ANCOVA), deneysel gruptaki öğrenciler arasında konuşma öz-yeterliklerinde dikkate değer gelişmeler olduğunu ortaya koydu. Mülakatlar, ChatGPT ile etkileşime girerek öğrencilerin konuşma becerilerinde özgüven geliştirmelerine, stres seviyelerini yönetmelerine ve öğrenme deneyiminden zevk almalarına yardımcı olduğunu ortaya koydu. Sonuçlar, ChatGPT gibi yapay zekâ uygulamalarının yabancı dil öğrencileri için uygun bir öğrenme ortamı oluşturarak konuşmada artan özgüvene yol açtığını gösterdi.

To Cite This Article: Yıldız, C. (2024). ChatGPT integration in EFL education: A path to improved speaking self-efficacy. *Novitas-ROYAL (Research on Youth and Language)*, *18*(2), 167–182. https://doi.org/10.5281/zenodo.13861137

1. Introduction

English as a foreign language (EFL) learners encounter significant obstacles when developing their speaking abilities. English is studied as a foreign language rather than naturally acquired in various contexts. Learners usually encounter the English language in educational environments within limited classroom hours. Language acquisition naturally becomes arbitrary and slow if not seen as necessary for surviving or satisfying needs or if not encountered daily (Crain, 1991). Additionally, learners of a foreign language may exhibit reluctance to engage in conversation in the target language due to insufficient practice, lack of confidence, or fear of making errors. Considering that students with a higher sense of self-efficacy are more likely to establish more challenging objectives, put in more effort to complete academic tasks, and utilize a range of adaptable and diverse learning methods, one possible solution to this problem is to increase students' confidence in their ability to speak EFL (Bandura & Schunk, 1981; Schunk, 1983).

Ever since Bandura (1997) introduced the notion of self-efficacy, scholars have been diligently examining the impact of self-efficacy on the learning process. Self-efficacy originates from the social cognition theory and thus pertains to an individual's perception of their competence in task completion based on their perceived abilities. It is a domain-specific construct (Bandura, 1986; 1997). In this regard, speaking self-efficacy pertains to learners' capacity to communicate effectively in the target language. A proficient learner feels self-assured and capable when facing challenges using the target language. Conversely, a low-efficacy learner feels inadequate when speaking the target language. Simply put, a person's confidence in their ability to complete a task can impact how well they perform it. Accordingly, studies have investigated the relationship between speaking self-efficacy and different factors. While various aspects of speaking self-efficacy have been recognized in research, it is essential to highlight that most studies investigate interpersonal interaction settings. However, little discussion has been conducted on how CALL affects speaking self-efficacy in EFL settings (Dong et al., 2022).

Tools powered by artificial intelligence (AI) offer numerous opportunities for English language learners to enhance their speaking self-efficacy. As mentioned, English language learning is a process in which many learners worldwide have a similar lack of exposure and application in speaking skills. These AI-based applications, such as ChatGPT, have the potential to enable learners to advance independently and enhance their self-assurance through the prompt provision of feedback and correction (El Shazly, 2021). Therefore, it is crucial to investigate how AI technology can potentially impact L2 learners' speaking self-efficacy. This study investigates the effect of artificial intelligence on language learners' confidence level and their capacity to speak English in higher education.

1.1. Artificial Intelligence in Language Education

Computational and information processing advances have enabled the accelerated development of artificial intelligence and its applications (Duan et al., 2019; Hwang et al., 2020; Tang et al., 2023). Hence, researchers and educators have developed AI-driven educational resources that have been employed in education to ameliorate the teaching and learning processes (Hwang et al., 2020; Wang et al., 2023). Additionally, artificial intelligence equips many advantages that can be noted in helping students better perceive their areas of strength and weakness in learning while using more personalized learning processes, enabling closer monitoring of students' development while offering more immediate and timely intervention, and enhancing the quality of education process through improved access to data (Göksel & Bozkurt, 2019).

Education, especially in language acquisition, calls for substantial time and dedication. AI technologies impact language teaching by providing a more advanced and efficient option than traditional models (Ermağan & Ermağan, 2022). These technologies bring about a new age in language learning. The era is characterized by interaction with technology, leading through natural language as if the systems were human beings. Moreover, from a behaviorist perspective, artificial intelligence and instruction can direct the learner at a self-administered pace, provide immediate corrective feedback, or reinforce the reward for an incorrect reply (Skinner, 1953). AI technologies have a great potential for improving speaking proficiency by personalized intervention and supporting speaking self-efficacy.

The key examples of AI language learning include personalized language materials and resources, AI chatbots, and AI language learning software (Rad, 2024). The effect of AI tools on speaking skills has been widely examined across different language learning environments. For instance, Li et al. (2020) investigated the effect of an AI-driven speech recognition system on improving pronunciation. People who received feedback from the AI system outperformed those who did not receive such feedback during their practice. Furthermore, multiple studies have been dedicated to exploring the effect of AI-based chatbots on language learning. Scholars underlined the opportunity for instant and interactive practice in AI chatbots' language. The studies showed that AI-based chatbots could significantly boost learners' motivation, engagement (Yangın Ersanlı, 2023; Zhang et al., 2019), WTC (Kim & Su, 2024), and communication confidence (Dizon, 2017; Sandeep, 2019). The relevant studies suggest that AI technologies, such as speech recognition systems and virtual tutors, can improve English language learners' speaking performance (Çakmak, 2022).

OpenAI created ChatGPT as a modern chatbot implementation with the ability to respond coherently and contextually. It utilizes much text data to create text in diverse tones and manners. ChatGPT is an excellent solution for generating material and can be used for various purposes. ChatGPT is an auspicious approach to automatically generating text that could eventually be applied to many purposes across diverse applications (Wu et al., 2023). ChatGPT can boost learners' confidence and smash away language difficulties to create a low-pressure atmosphere in which to practice overall language capability. ChatGPT can be used as a conversational partner when provided with proper prompts, and it can also give feedback about language usage, such as grammatical mistakes, vocabulary usage, and pronunciation errors. However, the application of ChatGPT in this context has not been investigated. Considering the lack of research in this area, investigating its potential in language education is required due to its vast ability to generate English content that might benefit EFL learners.

1.2. Speaking Self-Efficacy

According to social cognitive theory and comprehensive groundwork research, the concept of self-efficacy was posited by Bandura and his colleagues, and it refers to a person's belief in their ability to complete a task through their own efforts and skills (Bandura, 1997). This psychological construct, increasingly acknowledged, focuses on how individuals' beliefs in their ability to influence the environment dictate their actions to achieve desired results (Stajkovic & Luthans, 1998).

The four sources of self-efficacy are mastery experience (ME), vicarious experience (VE), social persuasion (SP), and physiological and emotional states (PES), according to social cognitive theory. Bandura (1997) characterized the first-hand experience as "enactive mastery experience"(p. 80). An individual's self-efficacy will be high if they have experience with the task and a strong personal connection. The second powerful influencer is vicarious experience.

Vicarious experience is acquiring knowledge or skills by observing others as they perform a task. Social persuasion encompasses positive verbal and nonverbal assessments received from others. Finally, Bandura (1997) explains that our mental and emotional states are the affective sources of self-efficacy. Following Bandura's (1997) theory, a positive experience increases self-efficacy, commitment to subsequent tasks, and the probability of success. According to Bandura (1986, 1997), self-efficacy varies depending on the task at hand, and researchers derive the most significant benefit from self-efficacy by directing their attention toward a particular context and domain of activity. Speaking self-efficacy describes a learner's capacity for appropriate target language communication. An effective learner is self-assured and competent enough to experience difficulty when utilizing the target language.

Speaking self-efficacy in online contexts has been the subject of some research in foreign language learning due to the quick development of digital media and communication technologies. Researchers have studied how computer-assisted language learning can increase the development of learners' speaking self-efficacy in different online settings. Zheng et al. (2009) explored the effect of chatting in Quest Atlantis (QA), a multi-user virtual environment (MUVE), on speaking self-efficacy and attitude toward learning the English language with Chinese EFL learners. QA students reported a higher level of speaking self-efficacy and felt more comfortable expressing their ideas in English than non-QA group learners. An investigation was conducted by Rahayu and Jacobson (2012) that examined the effect of MUVEs on Indonesian learners' speaking self-efficacy. Learners' speaking self-efficacy in the MUVE-based English program improved through continuous interaction with native speakers and the non-judgmental responses they received. Bárkányi (2021) found that the massive open online courses (i.e., LMOOC) participants showed more self-efficacy at the end of the course than at the beginning because the course environment in LMOOC was less threatening and less discouraging. Most researchers have studied the impact of different online contexts on learners' speaking self-efficacy. However, none of the EFL studies has explored the effect of AI communication tools on learners' speaking self-efficacy (Lee & Drajati, 2020). Thus, this study aims to investigate how ChatGPT, one of the most popular AI tools, affects language learners' speaking self-efficacy in English. Two factors were taken into consideration during the process of selecting the chatbot. Prior research has emphasized the significance of utilizing meticulously crafted and appropriate materials that align with the learners' levels and needs (Yang et al., 2022).

As a further point of interest, some researchers in the area of computer science evaluated the suitability of ChatGPT-generated materials for dialogue practice in language learning, and they confirmed that the majority of the materials that ChatGPT generated are best suited for students who are at the CEFR level A2 based on the Flesch Ease Reading scores (Fryer et al., 2019; Young & Shishido, 2023). The dialogues that were produced are suitable for students who have a proficiency level of CEFR A2 because they are able to comprehend the majority of the words that are used. Based on these factors, OpenAI's ChatGPT was determined to be the most appropriate technology for this research. The two research questions presented below served as the basis for this study.

1. To what extent does the use of ChatGPT improve the speaking self-efficacy of EFL learners?
2. What is the perception of EFL learners regarding using ChatGPT in speaking instruction?

2. Method

2.1. Research Design

The study's primary goal is to examine the effect of ChatGPT on the speaking self-efficacy level of English as foreign language learners. A mixed-method approach was used to achieve this goal. Specifically, data collection and analysis were done using quantitative and qualitative methods. The methodology employed for this study will be a sequential explanatory model. The sequential explanatory model comprises two distinct phases: a quantitative phase followed by a qualitative section (Creswell, 2003). In this model, a researcher initially gathers and examines the numerical data. Qualitative data is gathered and examined in a subsequent stage, aiding the researcher in elucidating the quantitative findings acquired in the initial phase. The reason for adopting this approach is that the numerical data and subsequent analysis comprehensively comprehend the research issue. The qualitative data and their analysis enhance and elucidate the statistical results by delving into participants' perspectives in greater detail (Creswell, 2003).

2.2. Participants

Two different sections of a compulsory general English course were used to carry out this project. Two classes were selected using a convenience sample method following official approval for conducting the research. The research study involved 65 college students studying EFL. They were all enrolled in the same university in Türkiye and took the general English course as part of their curriculum. The students were randomly divided into the control group (n=31) and the experimental group (n=34), which used the ChatGPT application. Both sections were taught by the same instructor and utilized the same course book, "SmartChoice 2" by Oxford University Press. The average age of these students was twenty years old, and they were enrolled in various majors, such as econometrics, economics, public administration, business administration, finance, and social work. Twenty-seven males and thirty-eight females made up the group. A placement exam administered at the beginning of the semester determined that their English proficiency was A2. The participants only used English in the English classes and had difficulty speaking English, potentially due to low self-efficacy, inadequate English proficiency, and specific personality traits (Tai & Chen, 2023). The instructor created the learning activities with ChatGPT and encouraged them to use it for extracurricular learning outside of the classroom to empower learners' speaking self-efficacy in English. The researcher selected ten students from among those who had completed the questionnaire to participate in interviews for the study's qualitative section, which will provide more information about the impact of ChatGPT activities on the speaking self-efficacy of EFL learners. The interview participants comprised six females and four males. The interview participants have majors in finance (three), economics (three), social work (two) and business administration (two).

2.3. Data Collection

The research tools used in the study were semi-structured interviews, speaking self-efficacy surveys, and the ChatGPT chatbot.

2.3.1. Speaking self-efficacy scale

For data collection, the EFL speaking self-efficacy scale (Wang & Sun, 2024) was utilized as the measuring instrument before and after ChatGPT activities. The scale developed by Wang and Sun (2024) was selected due to its specific focus on measuring speaking self-efficacy,

which fills a void in EFL research where there has been inconsistent use of general and task-specific self-efficacy measures. This scale enables a targeted evaluation of the distinct difficulties in cultivating speaking abilities. This scale consists of four sub-dimensions: linguistic self-efficacy with five items, self-regulatory self-efficacy with three items, delivery self-efficacy with three items, and performance self-efficacy with four items. After obtaining formal approval from the university, EFL students in selected classes were informed about the study's purpose and allowed to fill out questionnaires. They were also told that the information they submitted would be kept private. Participants were required to rate all items on a 7-point Likert scale.

2.3.2. ChatGPT as a learning tool

In this study, the effectiveness of ChatGPT on learners' speaking self-efficacy was investigated. A chatbot system that works for language learning usually comprises several necessary parts. Most frequently, those components will be a speech recognition module, audio content, and reference dialogue content (Young & Shishido, 2023). Ten objective tasks derived from the course's curriculum are presented in Table 1. Participants were expected to interact with ChatGPT based on the weekly discussion topics. Using the correct prompts, ChatGPT can turn into a conversation partner who can also give feedback about language use. Thus, learners in this study were provided with the prompt they should use to start a conversation with ChatGPT.

Participants used the same prompt each week only by changing the topic. The prompt used by learners was "Serve as my conversation partner and English teacher by posing interesting questions about (the topic of the week). Please only ask one question at a time, then give me time to respond. After I respond, ask different questions related to the subject. Do not limit yourself to just one line of questioning. Make sure the discussion is lively and engaging. Please correct my English grammar, phrasing, pronunciation, and vocabulary as we go along. If you understand, respond with "I understand" and move on to the next question." The instructor promoted the utilization of ChatGPT by learners as homework, providing explicit directions regarding its frequency and implementation. Student engagement was monitored via the shared chat feature of ChatGPT. Before being utilized in the classroom, ChatGPT was tested on various iOS and Android devices to ensure students could use it on their tablets and smartphones.

Table 1.
A sourcebook-based outline for both groups

Week	Topic
1	Asking somebody about his/her vacation
2	Making plans related to movies and music
3	Asking somebody about his/her experience in extreme sports
4	Talking about where you want to go and why
5	Talking about plans for a celebration
6	Refusing an invitation with a reason
7	Asking about a recent shopping trip
8	Talking about applying for a vacation job
9	Talking about an accident or injury you had
10	Talking about what you used to do when you were younger

2.3.3. Semi-structured interview

In order to gain experience-based insights into the impact of ChatGPT activities on the speaking self-efficacy of EFL learners, ten learners from the experimental group participated in a semi-

structured interview. The interview was conducted voluntarily. Two questions were addressed to those volunteers: 1) What are your thoughts regarding using ChatGPT to enhance English speaking skills? 2) To what extent do you believe the ChatGPT activities have improved your speaking self-efficacy in English? What are the reasons, if any? The interviews were conducted in Turkish to create a comfortable environment for the participants and to encourage them to provide more detailed accounts of their experiences. The duration of all interviews ranged from 15 to 25 minutes.

2.4. The ChatGPT infusion procedure

Several stages were included in this research project. The control and experimental groups were randomly chosen from two classes. Both groups first filled out a speaking self-efficacy questionnaire before the intervention. The same teacher instructed both the control group and the experimental group to get rid of any potential bias. Approximately ten weeks were spent experimenting. Both groups attended two sessions lasting ninety minutes each week.

Three stages are included in the English language classes: content presentation, practice, and production. When it came to the content presentation stage, all the students utilized the standard lecture materials in the textbook to acquire the necessary vocabulary, expressions, and grammar for each unit. The instructor was the primary source of information. During the practice stage, they carried out listening activities and completed exercises such as fill-in-the-blanks, sentence transformations, and the creation of sentences utilizing particular grammar structures. During the production phase, learners participated in a communication task through either pairs or groups. The experimental group used ChatGPT as homework outside the class. Before starting to use the ChatGPT, users need to complete several steps. These steps include downloading and installing the ChatGPT on any of their devices, creating an account, and adjusting the profile.

Further on, they investigated the ChatGPT's characteristics and resources. Hence, with ChatGPT, users can go through a speaking procedure by recording their speeches. They can talk about numerous topics. ChatGPT uses innovative technology and advanced, automatic speech recognition to evaluate users' speech. It provides instant feedback on users' speech concerning English grammar, phrasing, and word choice. Users can evaluate their speech based on this feedback. Considering these steps and features, users could integrate this application into their language-learning process. At the end of each week, learners shared the transcripts of the recorded text with their instructor through email. Thus, the instructor would follow students' improvements.

2.5. Data Analysis

Descriptive and inferential statistical methods were employed to analyze the data. Firstly, the means and standard deviations of the data were computed to evaluate the speaking self-efficacy levels in each group. Then, a statistical analysis was conducted to determine the questionnaire's reliability. The internal reliability of the scale was determined to be 0.74 for the pre-test. This value was considered satisfactory regarding the scale's consistency and accuracy. Then, an independent-sample t-test was conducted to compare the speaking self-efficacy levels of the experimental and control groups before the intervention to confirm the groups' comparability before the intervention. In the end, the impact of the treatment was examined using an ANCOVA test after conducting initial analyses on the SPSS. This analysis used learners' scores on the pre-intervention administration of the scale as the covariate. To account for potential pre-existing differences between the experimental and control groups,

ANCOVA was chosen instead of a t-test or ANOVA. Pre-test results are used as a covariate in ANCOVA, which corrects for initial disparities differently than a t-test or ANOVA. This allows for a more effective separation of the impact of the treatment and a clearer picture of its effect. Preliminary checks were conducted to ensure no violation of the assumptions of normality, linearity, or homogeneity of variances. The internal reliability of the scale was found to be 0.95 after the intervention, and Levene's test of equality of error variances table yielded a p-value greater than 0.05.

The researcher transcribed the interviews via an audio recorder for qualitative data analysis. Subsequently, digressions and repetitions were removed to enhance the clarity of the transcript. The interviewees' responses were subsequently condensed into more concise formulations using the meaning condensation method (Kvale, 1996). As a result, lengthy interview transcripts were condensed into concise formulations. The transcripts were divided into meaningful segments (Merriam, 1998). The segments comprised pre-established interview themes, including students' perceptions of their speaking self-efficacy, the impact of ChatGPT on their speaking self-efficacy, and the use of ChatGPT for improving speaking abilities. The researcher formulated assertions by comparing the responses of the interviewees. Direct interview quotes also substantiated these claims (Erickson, 1986). The qualitative data coding process involves two coders with expertise in English language education to ensure greater consistency and validity of the results. The reliability was ensured through a substantial inter-coder agreement rate of 88%.

3. Results

3.1. Quantitative phase

3.1.1. Descriptive statistics for self-efficacy

The main goal of the first research question was to investigate the impact of ChatGPT on EFL learners' levels of speaking self-efficacy. As a first step, the participants were requested to rate their perceptions of their speaking self-efficacy in English on a seven-point scale ("strongly disagree") to ("strongly agree"). The experimental (M=3.86, SD=0.29) and control (M=3.46, SD=0.49) groups reported low speaking self-efficacy in English. They struggled to strongly support any of the items in the questionnaire, as indicated in Table 2. The participants in the experimental group demonstrated diminished levels of speaking self-efficacy related to their ability to speak with correct pronunciation, intonation, and liaison (M=3.17, SD=0.86) and speak with few pauses or filters (M=3.02, SD=0.71). On the other hand, the participants in the control group reported lower scores for controlling their stress levels when speaking English in the classroom (M=3.35, SD=0.83) and thinking of goals before speaking (M=3.35, SD=0.83). The scale assessed learners' speaking self-efficacy in four domains: linguistic self-efficacy (LSE), self-regulatory efficacy (SRE), delivery self-efficacy (DSE), and performance self-efficacy (PSE). The experimental group reported the lowest level of speaking self-efficacy in the linguistic self-efficacy domain (M=3.52, SD=0.48). In contrast, the lowest mean score was found in the performance self-efficacy domain for the control group (M=3.16, SD=0.82).

The second set of surveys was executed ten weeks after administering the initial EFL-speaking self-efficacy surveys. A total of ten different speaking tasks involving ChatGPT were carried out by the experimental group over the ten weeks, as detailed in Table 1. Following the data presented in Table 2, there were distinctions between the two groups. The speaking self-efficacy levels of the experimental group were significantly more favorable than those of the control group (EG=5.76, CG=4.04). The experimental group reported that

they could speak fluently (M=6.02 SD=0.57) with grammatical accuracy (M=5.94, SD=0.60), and they could understand the most challenging material presented in the speaking course (M=5.64, SD=0.64). They also indicated that they could speak confidently (M=5.88, SD=0.68) and enjoy speaking English outside the classroom (M=6.02, SD=0.62).

Unlike the experimental group participants, the control group participants' speaking self-efficacy levels slightly increased in the mean, but it was still within the "uncertain" range. These participants reported the lowest mean score for the self-regulatory efficacy domain. Consequently, these findings suggest that the utilization of ChatGPT impacted the learners' speaking self-efficacy in English. A statistical analysis was performed to verify the accuracy of this claim.

Table 2.

Participants' level of speaking self-efficacy in English

Speaking Self-Efficacy Dimensions	Before intervention				After intervention			
	Experimental group		Control group		Experimental group		Control group	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Linguistic Self-efficacy (LSE)	3.52	0.48	3.70	0.53	5.58	0.29	3.98	0.41
Self-Regulatory Efficacy (SRE)	3.75	0.53	3.51	0.52	5.64	0.44	3.96	0.55
Delivery Self-efficacy (DSE)	3.77	0.51	3.52	0.55	5.93	0.35	4.01	0.66
Performance Self-Efficacy (PSE)	3.86	0.45	3.46	0.49	5.76	0.33	4.04	0.48
Overall	3.70	0.29	3.56	0.41	5.71	0.19	4.00	0.39

3.1.2. Analysis of statistical differences between groups

Initially, an independent-sample t-test was conducted to compare the speaking self-efficacy levels of the experimental and control groups before the intervention. There was no significant difference in speaking self-efficacy levels for the experimental group (M=3.70, SD=0.29) and control group (M= .56, SD=0.41); t(65)=-1.552, p=.127), which shows that the two groups were relatively similar and homogeneous, as Table 3 shows.

Table 3.

Independent samples t-test of speaking self-efficacy before intervention

	95% Confidence Interval of the Difference							
	Mean	SD	Lower	Upper	t	df	Sig. (2-tailed)	Cohen's d
Experimental Group	3.70	0.29	- .320	.040	-1.552	54.41	.127	0.394
Control Group	3.56	0.41						

The impact of the treatment was examined using an ANCOVA test after conducting initial analyses on the SPSS. Table 4 displays the outcomes of the ANCOVA test (F(1, 65)= 491.184, p=.000), indicating significant differences in speaking self-efficacy between the experimental and control groups after the intervention. These findings indicate that learners' speaking self-efficacy levels increased significantly due to the ChatGPT intervention. More specifically, the experimental group's speaking self-efficacy levels were higher than the control group's, as indicated by the adjusted post-test means. This shows that the learners' confidence and perceived English-speaking ability were raised using ChatGPT as an intervention tool. The ANCOVA test's significant difference highlights the treatment's effectiveness and potential advantages in language learning environments. Therefore, the

ChatGPT intervention had a significant impact on increasing the learner's speaking self-efficacy levels.

Table 4.
ANCOVA results for the post-intervention speaking self-efficacy

Dimension	Group	Source	*SS*	*df*	*MS*	*F*	*p*
Speaking self-efficacy	Experimental Control	Group Total	46.585 1613.076	1 65	46.585	491.184	.000

a. R2 = .890 (Adjusted R2 = .886)

3.2. Qualitative Data Analysis

The second research question focused on investigating the participants' viewpoints regarding using ChatGPT to improve speaking skills in English. This question inquired about the participants' communication experiences, their speaking self-efficacy in English, the potential future of using ChatGPT for improving communicative skills, and the advantages and disadvantages of utilizing this technology in language learning. During the semi-structured interview, ten students voluntarily participated. The data from the interview was initially transcribed in Turkish, and then it was translated into English. During the analysis of the interview data, five significant themes concerning the utilization of ChatGPT for facilitating English language learning were consistently derived. These themes concern increasing confidence in speaking, improving language proficiency, enjoyment, convenience, and limited feedback on pronunciation. Table 5 contains the thematic codes that were derived from the interview data.

Table 5.
Themes from Interviews on Learners' Perspectives of ChatGPT Usage

Category	Code Description
Increasing confidence in speaking	This theme pertains to learners experiencing enhanced self-assurance in their oral communication skills through using ChatGPT. It includes emotions of confidence, courage, and faith in one's capacity to communicate in English proficiently. (Percentage: 90%)
Improving language proficiency	This theme represents the concept of learners improving their comprehensive language abilities, encompassing vocabulary, grammar, and fluency, by engaging with ChatGPT. It demonstrates the growth of proficiency and expertise in utilizing the English language. (Percentage: 80%)
Enjoyment	This theme exemplifies the concept of learners deriving enjoyment and contentment from utilizing ChatGPT for oral exercises. It encompasses emotions of pleasure, amusement, and involvement throughout learning a new language. (Percentage: 85%)
Convenience	This theme represents the concept of learners appreciating the convenience and availability of ChatGPT for speaking practice, which is represented by this code. It provides the flexibility to practice whenever and wherever it is most convenient. (Percentage: 75%)
Limited feedback on pronunciation	This code represents the concept of learners receiving feedback on their pronunciation from ChatGPT but perceiving that the feedback is insufficient in addressing different aspects of pronunciation. It includes instances in which students have expressed the opinion that the feedback provided by ChatGPT regarding pronunciation could be more specific or helpful. (Percentage: 85%)

Most of the learners reported that their confidence in their ability to speak English improved after they had practiced with ChatGPT. When practicing with ChatGPT, they reported that

they felt more relaxed and did not experience any anxiety about making mistakes. For instance, two students, who typically feel more comfortable remaining silent during speaking class, reported that they had observed some changes, such as decreased anxiety regarding the possibility of making mistakes and increased engagement:

> *I usually keep quiet in speaking classes because I'm afraid my friends will laugh at me if I say something wrong. When I speak with ChatGPT, I don't have such fears or worries. I can speak whatever I want without fear.*

> *Actually, speaking English is not something to be afraid of. I had overestimated it before. For some reason, I used to get very nervous and anxious when speaking English in a classroom environment, but with ChatGPT, I don't feel that way. I can try again and again.*

Also, learners indicated that the direct feedback they received from ChatGPT helped them increase their confidence in their speaking abilities:

> *I had the opportunity to discuss many topics using the ChatGPT robot, and it corrected my mistakes directly and helped me understand instantly. I think my confidence has improved.*

Interacting with ChatGPT improved their language proficiency. Learners indicated that ChatGPT immediately pointed out their vocabulary and grammar mistakes and presented their correct usages:

> *In my opinion, one of the best things about ChatGPT is that it corrects what I say wrong. For example, I always forget to add s to verbs when I speak. ChatGPT corrected me every time.*

> *While talking with ChatGPT, I can directly ask for the words I don't know. I don't need to search. The most difficult point for me while talking was my lack of vocabulary. I think it helped me a lot with vocabulary.*

Most participants enjoyed using ChatGPT to learn English compared to traditional classroom activities. They indicated that they had the freedom to direct conversations with ChatGPT:

> *ChatGPT was much fun to talk to. We discussed the topic weekly, and the conversation was very enjoyable in every way. Talking to ChatGPT is like chatting with a close friend. I never got bored.*

> *It's like talking to a native speaker. I really enjoyed it. The book's activities bored and limited me a lot. It was like a forced conversation.*

Convenience was another important theme. Learners indicated that they could find the opportunity to practice speaking English outside the class at any time:

> *I installed ChatGPT on my phone, and I can practice English anywhere. I think it's awesome. I don't remember speaking English at home before.*

Concerning the negative side of ChatGPT, most learners pointed out its limited capacity to provide feedback on pronunciation. Most of the feedback they received during the conversation was primarily based on grammar and vocabulary. ChatGPT did not give as much feedback on pronunciation as on vocabulary or grammar mistakes. They reported not getting any feedback on intonation or stress patterns:

> *Most of the feedback ChatGPT gave me while speaking was about grammar or vocabulary explanations. When I wanted to know the pronunciation of a word, I asked. I have problems saying words correctly. I wish it could correct my pronunciation during the conversation.*

All in all, ChatGPT positively impacted the psychological and linguistic components of foreign language learning. ChatGPT offers an encouraging and engaging learning environment, allowing them to practice speaking English, ultimately increasing self-confidence and proficiency. Those who took part in the chatbot exercises were able to improve their vocabulary and grammar skills even though they expressed a desire for more detailed feedback on pronunciation.

4. Discussion

This study aimed to investigate the impact of ChatGPT on the speaking self-efficacy of EFL learners. The study employed both quantitative and qualitative methods to achieve this goal. Based on the findings gained from the first questionnaire, the initial situation can be characterized by low speaking self-efficacy in the English classroom. Both groups reported that they struggled with speaking in English in many ways, such as correct grammar usage, proper fluency and pronunciation, and even control over their own personal stress matter during the conversation. The results of the second questionnaire showed differences between the control and experimental groups. The experimental group was found to have higher confidence in speaking and stress control, considerably more oral fluency, and more expression of the fun they had conversing. Therefore, the results showed that the speaking self-efficacy levels of EFL learners were positively affected by ChatGPT.

Statistical analyses nevertheless revealed differences with all parameters. Specifically, the experimental group felt more confident in speaking without stress. Utilizing ChatGPT increased the learners' feelings of better preparedness for more complex speaking tasks and control over the process. This finding was supported by recent studies that reported that AI-based speaking activities are fun to study, relax anxiety, and help learners gain confidence in their speaking abilities (e.g., Kim & Su, 2024; Rad, 2024; Tai & Chen, 2023; Zou et al., 2023).

For the second research question, the data collected during the interviews demonstrate that the learning use of ChatGPT positively influenced the learners' affective and cognitive aspects. To be more precise, the majority of the interviewees reported that the way of interacting with ChatGPT boosted their level of self-confidence and ability to control anxiety during a conversation. In addition, the group agreed on the statement that their ability to structure the words meaningfully had improved in terms of grammatical accuracy due to the immediate feedback from the ChatGPT, despite some problems related to the pronunciation feedback. Learners found the learning process enjoyable due to the flexible and convenient features of ChatGPT. The findings are consistent with the previous research, which demonstrated the effectiveness of AI chatbots in enhancing the oral English proficiency of learners (Rad, 2024) and increasing engagement, comfort, confidence, and motivation (Tai & Chen, 2023).

The results of our study have several implications for educational settings. Artificial intelligence chatbots can be incorporated into the classroom setting of an EFL course as a practical method for helping learners improve their speaking self-efficacy and, ultimately, making it easier for them to communicate in English. Previous research has found that mastery experience is the most potent factor in determining one's level of speaking self-efficacy across the four dimensions of speaking self-efficacy (Mills et al., 2007; Wang & Sun, 2024). The self-confidence of EFL learners in adopting metacognitive strategies, delivering self-controlled speaking

presentations, and appraising their speaking competence is influenced by their prior successful speaking experiences. Language educators and practitioners can utilize artificial intelligence applications to create a more meaningful, captivating, and less daunting atmosphere for practicing English, facilitating the cultivation of significant experiences and self-assurance in English. Positive experiences in an AI-driven setting can help learners become more confident in their ability to perform similar tasks in the future, raising their speaking self-efficacy. This is particularly encouraged in an EFL environment, where learners have limited opportunities to speak English outside the classroom.

These applications' language learning programs are tailored to meet each student's specific requirements and interests. Personalized content is created by them, which boosts learner engagement and motivation, ultimately increasing learners' confidence in their ability to communicate effectively (Rad, 2024). Students have the opportunity to improve their speaking abilities by employing continuous interaction with AI chatbots, as well as receiving instantaneous recognition for their efforts. In addition, these applications offer a low-risk and non-critical setting, making them ideal for practising and receiving feedback. As a result, this helps reduce the anxiety learners experience and boosts their confidence.

5. Conclusion

This study examined the influence of ChatGPT on the speaking self-confidence of EFL learners. At first, both groups had low confidence in their speaking ability and faced difficulties with grammar, fluency, pronunciation, and managing stress during conversations. Following the intervention, the experimental group demonstrated elevated levels of self-assurance in verbal communication, enhanced ability to manage stress, improved proficiency in speaking, and increased pleasure in engaging in conversations compared to the control group. Therefore, the study determined that ChatGPT positively impacted the levels of speaking self-efficacy among EFL learners.

Although the study has shown positive results, several limitations should be discussed. First, the sample was relatively small. A larger sample would have a better representation of results and would be more widely applicable. Second, the study was conducted in a relatively short period of ten weeks. It might also not be long enough to comprehensively analyze the long-term effects of ChatGPT on speaking self-efficacy. Finally, this research was conducted using a quasi-experimental method, meaning no extraneous variables that could influence the levels of speaking self-efficacy were controlled.

This study's results have various practical implications for decision-makers, educators, and researchers. Language teachers should consider incorporating AI chatbots into their curriculum to offer students supplementary opportunities to practice speaking in a low-stress setting. ChatGPT can assist students in overcoming anxiety and fostering confidence in their speaking abilities, thereby enhancing their overall language proficiency. Educational institutions should support the adoption of AI technologies in language learning programs. Providing training and resources for teachers to use these tools effectively can create a more engaging and supportive learning environment for students. This study demonstrates how AI chatbots, such as ChatGPT, can help EFL learners feel more confident when speaking.

Future research should concentrate on several essential areas to build on the results of this study. First, larger sample sizes should be employed to strengthen the results' generalizability and offer a more accurate depiction of the effect of AI chatbots like ChatGPT on speaking self-efficacy. A ten-week intervention may not fully capture the long-term effects of the

technology, so longitudinal studies are also required to evaluate ChatGPT's effects on learners' speaking abilities and self-confidence. Furthermore, more rigorous experimental designs that account for unrelated variables will support the validity of the results and guarantee that the effects observed are actually related to the intervention. To ascertain the efficacy of AI chatbots in diverse educational contexts, researchers ought to consider a range of populations, encompassing age groups, proficiency levels, and cultural backgrounds. Lastly, investigating ChatGPT's integration with other instructional strategies and technologies may shed light on how to design thorough and efficient language learning curricula. These new lines of inquiry will contribute to establishing AI's place in improving language instruction.

Note on Ethical Issues

The author confirms that ethical approval was obtained from Uşak University on 09/05/2024 with the decision number 2024-118. It should be noted that ChatGPT was only used for educational and research purposes in this study. No text was generated using ChatGPT.

Conflict of Interest

The authors declare no conflict of interest.

References

Bandura, A. (1986). *Social foundations of thought and action*: *A social cognitive theory*. USA: Prentice-Hall.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.

Bandura, A., & Schunk, D. H. (1981). Cultivating competence, self-efficacy, and intrinsic interest through proximal self-motivation. *Journal of personality and social psychology, 41*(3), 586-598. https://doi.org/10.1037/0022-3514.41.3.586

Bárkányi, Z. (2021). Motivation, self-efficacy beliefs, and speaking anxiety in language eMOOCs. *ReCALL*, *33*(2), 143–160. https://doi.org/10.1017/S0958344021000033

Crain, S. (1991). Language acquisition in the absence of experience. *Behavioral and Brain Sciences, 14*(4), 59–612. https://doi.org/10.1017/S0140525X00071491

Creswell, J. W. (2013). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). London: Sage.

Çakmak, F. (2022). Chatbot-human interaction and its effects on EFL students' L2 speaking performance and speaking anxiety. *Novitas-ROYAL (Research on Youth and Language), 16(*2), 113–131.

Dizon, G. (2017). Using intelligent personal assistants for second language learning: A case study of Alexa. *TESOL Journal, 8*(4), 811–830. https://doi.org/10.1002/tesj.353

Dong, L., Jamal Mohammed, S., Ahmed Abdel-Al Ibrahim, K., & Rezai, A. (2022). Fostering EFL learners' motivation, anxiety, and self-efficacy through computer-assisted language learning- and mobile-assisted language learning-based instructions. *Frontiers in Psychology, 13*, 1–15. https://doi.org/10.3389/fpsyg.2022.899557

Duan, Y., Edwards, J. S., & Dwivedi, Y. K. (2019). Artificial intelligence for decision making in the era of Big Data–evolution, challenges, and research agenda. *International journal of information management*, *48*, 63–71. https://doi.org/10.1016/j.ijinfomgt.2019.01.021

Dwyer, K. K., & Fus, D. A. (2002). Perceptions of communication competence, self-efficacy, and trait communication apprehension: Is there an impact on basic course success? *Communication Research Reports, 19*(1), 29–37. https://doi.org/10.1080/08824090209384829

El Shazly, R. (2021). Effects of artificial intelligence on English speaking anxiety and speaking performance: A case study. *Expert Systems*, *38*(3), e12667. https://doi.org/10.1111/exsy.12667

Erickson, F. (1986). Qualitative methods in research on teaching.In C. M. Wittrock (Ed.), *Handbook of research on teaching* (pp. 145–152) New York, USA: Macmillan.

Ermağan, E., & Ermağan, İ. (2022). Innovative technology and education: Artificial intelligence and language learning in Turkey. *Shanlax International Journal of Education, 11*(S1), 201–209. https://doi.org/10.34293/education.v11iS1-Dec.6085

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

Göksel, N. & Bozkurt, A. (2019). Artificial intelligence in education: Current Insights and future perspectives. In S. Sisman-Ugur & G. Kurubacak (Eds.), *Handbook of Research on learning in the age of transhumanism* (pp. 224–236). Hershey, PA: IGI Global. https://doi.org/10.4018/978-1-5225-8431-5.ch014

Hwang, G. J., Xie, H., Wah, B. W., & Gašević, D. (2020). Vision, Challenges, roles and research issues of artificial intelligence in education. *Computers and Education: Artificial Intelligence, 1*, 100001. https://doi.org/10.1016/j.caeai.2020.100001

Kim, A., & Su, Y. (2024). How implementing an AI chatbot impacts Korean as a foreign language learners' willingness to communicate in Korean. *System,* 103256. https://doi.org/10.1016/j.system.2024.103256

Kvale, S. (1996). *Interviews: An introduction to qualitative research interviewing*. Thousand Oaks, USA: Sage Publication.

Lee, J. S., & Drajati, N. A. (2020). Willingness to communicate in digital and non-digital EFL contexts: Scale development and psychometric testing. *Computer Assisted Language Learning*, *33*(7), 688–707. https://doi.org/10.1080/09588221.2019.1588330

Li, J., Xie, H., Li, D., & Lu, X. (2020). Effects of an AI-based speech recognition system on improving EFL learners' pronunciation. *Computers & Education, 156*, 103952.

Merriam, S. B. (1998). *Qualitative research and case study applications in education.* San Francisco: Jossey-Bass Publishers.

Mills, N., Pajares, F., & Herron, C. (2007). Self-efficacy of college intermediate French students: Relation to achievement and motivation. *Language learning, 57*(3), 417–442. https://doi.org/10.1111/j.1467-9922.2007.00421.x

Rad, H. S. (2024). Revolutionizing L2 speaking proficiency, willingness to communicate, and perceptions through artificial intelligence: A case of Speeko application. *Innovation in Language Learning and Teaching*, 1–16. https://doi.org/10.1080/17501229.2024.2309539

Rahayu, P., & Jacabson, M. J. (2012). Speaking self-efficacy and English as a foreign language: Learning processes in a multi-user virtual environment. In J. Peterson, O. Lee, T. Islam, & M. Piscionerri (Eds.), *Effectively Implementing information communication technology in higher education in the Asia-Pacific region* (pp. 161–182). USA: Nova Science Publishers.

Schunk, D. H. (1983). Ability versus effort attributional feedback: Differential effects on self-efficacy and achievement. *Journal of Educational Psychology, 75*(6), 848–856. https://doi.org/10.1037/0022-0663.75.6.848

Skinner, B. F. (1953). Some contributions of an experimental analysis of behavior to psychology as a whole. *American Psychologist, 8*(2), 69–78. https://doi.org/10.1037/h0054118

Tang, K. Y., Chang, C. Y., & Hwang, G. J. (2023). Trends in artificial intelligence-supported e-learning: A systematic review and co-citation network analysis (1998–2019). *Interactive Learning Environments*, *31*(4), 2134–2152. https://doi.org/10.1080/10494820.2021.1875001

Sandeep, S. S. (2019). Learning English language through Amazon Alexa for Indian students. *Journal of the Gujarat Research Society*, *21*(10), 619–622.

Stajkovic, A. D., & Luthans, F. (1998). Social cognitive theory and self-efficacy: Goin beyond traditional motivational and behavioral approaches. *Organizational dynamics*, *26*(4), 62–74. https://doi.org/10.1016/S0090-2616(98)90006-7

Tai, T. Y., & Chen, H. H. J. (2023). The impact of Google Assistant on adolescent EFL learners' willingness to communicate. *Interactive Learning Environments, 31*(3), 1485–1502. https://doi.org/10.1080/10494820.2020.1841801

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., & Sun, P. P. (2024). Development and validation of scales for speaking self-efficacy: Constructs, sources, and relations. *PlosOne*, *19*(1), e0297517. https://doi.org/10.1371/journal.pone.0297517

Wu, T., He, S., Liu, J., Sun, S., Liu, K., Han, Q. L., & Tang, Y. (2023). A brief overview of ChatGPT: The history, status quo and potential future development. *IEEE/CAA Journal of Automatica Sinica*, *10*(5), 1122–1136. https://doi.org/10.1109/JAS.2023.123618

Yang, H., Kim, H., Lee, J. H., and Shin, D. (2022). Implementation of an AI chatbot as an English conversation partner in EFL speaking classes. *ReCALL, 34*, 327–343. http://doi.org/10.1017/S0958344022000039

Yangın Ersanlı, C. (2023). The effect of using augmented reality with storytelling on young learners' vocabulary learning and retention. *Novitas-ROYAL (Research on Youth and Language), 17*(1), 62–72.

Young, J. C., & Shishido, M. (2023). Investigating OpenAI'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. https://doi.org/10.14569/IJACSA.2023.0140607

Zhang, D., Tang, J., Gao, L., & Wei, X. (2019). Chatbot-assisted language learning: An empirical study of learners' perceptions and attitudes. *IEEE Access, 7*, 13310–13319.

Zheng, D., Young, M. F., Brewer, R. A., & Wagner, M. (2009). Attitude and self-efficacy change: English language learning in virtual worlds. *CALICO Journal*, *27*(1), 205–231. https://doi.org/10.11139/cj.27.1.205-231

Zou, B., X. Guan, Y. Shao, and P. Chen. (2023). Supporting speaking practice by social network-based interaction in artificial intelligence (AI)-assisted language learning. *Sustainability, 15*(4), 2872. https://doi.org/10.3390/su15042872