

The Effect of ChatGPT as a Pen-Friend on Primary School Learners' Reading and Writing Skills and Their Attitudes Towards English Courses

Dilara Yücetürk

Amasya University, Institute of Science and Technology, Department of Computer Education and Instructional Technology, Amasya, Türkiye
ORCID: 0000-0002-3742-3985

Özgen Korkmaz*

Bandırma Onyedi Eylül University, Engineering and Natural Science Faculty, Department of Computer Engineering, Bandırma, Balıkesir, Türkiye
ORCID: 0000-0002-0880-1318

Mehmet Kara

Amasya University Education Faculty, Department of Educational Sciences, Amasya, Türkiye
ORCID: 0000-0003-2758-2015

Article history

Received:
23.01.2026

Received in revised form:
19.03.2026

Accepted:
01.04.2026

Key words:

ChatGPT; reading skills;
writing skills; elementary
students; AI in education;
language learning

In the literature, there is not enough evidence specifically regarding the contribution of ChatGPT to the English education of primary school students. This calls for further studies empirically demonstrating the effectiveness of generative Artificial Intelligence (AI) tools in language learning in primary school settings. This sort of study is also useful in providing implications for effectively integrating these novel tools into school settings. Therefore, this study focuses on the impact of ChatGPT on the English reading and writing skills of elementary school students. This study explores the impact of using ChatGPT as a pen pal on young learners' reading and writing skills in English courses and their attitudes towards language learning. Conducted with 55 elementary school students, the study employed a quasi-experimental design. The experimental group interacted with ChatGPT weekly to practice their reading and writing skills, while the control group followed traditional teaching methods. Pre-tests and post-tests were administered to assess students' progress in reading and writing, along with an attitude scale to measure changes in their perception of the English course. The results showed that both groups significantly improved their reading and writing skills, but no significant difference was observed between the groups. Although ChatGPT proved to be an engaging tool, its impact was comparable to traditional methods. This study highlights the potential of AI-driven tools like ChatGPT in supporting language learning, suggesting further research is needed to explore its long-term effects.

Introduction

As a result of the globalization and digitalization of our world, the use of English, accepted as the common language, is becoming increasingly widespread, and the effective utilization of this language is gaining growing importance (Dash & Gandhi, 2022). However,

* Correspondency: ozgenkorkmaz@gmail.com

English education faces challenges such as insufficient funding, limited access to quality language education, and unequal opportunities (Ergün, 2022). In the Turkish context of English education, despite long-standing integration of foreign language courses and various program changes to address evolving needs, there are persistent challenges. Türkiye faces difficulties in teaching foreign languages, evident in international proficiency rankings and national exams. Current research further identifies these struggles as a holistic failure rooted in methodological, linguistic, and systemic barriers; notably, the heavy reliance on grammar-based instruction and the pressure of high-stakes exams prioritize rote memorization over communicative proficiency (Aydın, 2025). These factors, combined with a lack of real-life exposure, create significant psychological barriers, such as speaking anxiety and low self-confidence among learners (Aydın, 2025). Moreover, the lack of sufficient environments for students to practice language skills poses a significant challenge. The absence of opportunities for practical application impedes effective language learning. Artificial Intelligence (AI) tools can contribute to the creation of interactive and immersive learning environments worldwide, providing personalized language exercises, simulations, and real-life scenarios to address the global lack of sufficient environments for students to practice language skills (Owan et al., 2023). This innovative approach aligns with the evolving landscape of educational technology and offers a solution to the universal problem of limited practical language practice (Conati & Lallé, 2023). Leveraging AI applications can help overcome the absence of opportunities for practical application, thereby facilitating effective language learning (Menekşe, 2023).

AI holds great promise for improving language learning by offering individualized and flexible learning opportunities. The capacity of AI systems to evaluate unique learning patterns, strengths, and shortcomings is one obvious benefit. These systems maximize the effectiveness of the educational process by customizing language lessons to each learner's specific needs using complex algorithms (Rugaiyah, 2023). Furthermore, as a highly flexible and personalized learning partner, ChatGPT provides instantaneous feedback tailored to individual linguistic needs, significantly enhancing the autonomy of the learner (Kafabih, 2026). AI-powered language platforms can also mimic real-world conversational situations, providing learners with an opportunity to practice and improve their language skills in a safe yet dynamic setting. By participating in real-world discussions, users can acquire practical language skills through this immersive approach, which enhances the effectiveness and enjoyment of the learning process (Kumar et al., 2023). Recent research further confirms that the integration of such AI tools into language pedagogy fosters a more interactive and accessible learning environment, bridging the gap between theoretical knowledge and practical application (Kafabih, 2026).

The promise of AI's further evolution lies in its potential application to language instruction (Rusmiyanto, 2023). ChatGPT, a product of OpenAI within the GPT-3.5 architecture, represents a significant advancement in the field of AI with promising applications in language teaching. It plays a crucial role in assisting students with AI-based language acquisition and has a lot of room to grow as long as technology advances (Kohnke, Moorhouse, & Zou, 2023). ChatGPT has undergone extensive training on a wide variety of internet content and is designed to comprehend and produce text that is similar to that of a human. This advanced conversational AI model seems very effective at understanding and reacting to different language-based activities. Because of its adaptability, it can actively participate in conversations, respond to inquiries, and offer knowledge on a variety of subjects (Wu et al., 2023). Within the field of language learning, ChatGPT stands out as a potent instrument, drawing on its vast expertise to enhance the entire educational process and provide students with useful direction on their language learning path (Bin-Hady, Al-Kadi, Hazaea, & Ali, 2023). One major factor in ChatGPT's efficacy is how well it comprehends



context and can produce well-reasoned responses. Because of this, ChatGPT is a priceless tool for students looking to improve their language abilities using AI-based learning environments (Hong, 2023). Through its sophisticated AI capabilities, ChatGPT plays a critical role in improving the overall quality of language learning, whether it be by facilitating engaging discussions, producing content, or providing guidance. The potential uses of ChatGPT in language instruction are probably going to grow as more advancements take place, making it a key component of the changing AI-driven education scene (Topsakal & Topsakal, 2022).

Students can effectively use ChatGPT to enhance their language skills in various ways. Firstly, ChatGPT can aid students in creating personalized language learning materials. By requesting content that aligns with their interests and requirements, students can experience a more customized and engaging language learning journey. Additionally, students can seek assistance with grammar topics, expand their vocabulary, and enhance different language skills by tapping into the extensive knowledge resources offered by ChatGPT. The model can serve as a valuable tool for language learners to receive personalized support tailored to their individual needs (Javaid, et al., 2023). Furthermore, engaging in interactive conversations with ChatGPT allows students to improve their abilities in natural language communication. ChatGPT provides an ideal environment for practicing conversational skills by answering questions and engaging in dialogue. However, recent research suggests that the integration of AI-generated feedback is not always seamlessly accepted by learners. Specifically, L2 students may exhibit resistance or rejection toward ChatGPT-generated feedback due to concerns regarding its perceived accuracy, pedagogical depth, or a lack of personal connection compared to human instructors (Chen et al., 2026). In conclusion, it offers students a range of tools to develop their language skills at their own pace and according to their learning styles (Adiguzel, Kaya, & Cansu, 2023), the effectiveness of these tools is significantly mediated by students' trust and their critical evaluation of AI-driven output.

AI models such as ChatGPT can improve language learning in higher education and vocational training by helping with writing, summarizing, research, translation, and analysis. It facilitates instructional tasks, essay writing, and assessment design in secondary school (Bekou, Mhamed, & Assissou, 2024). Just like every level of students, Students in elementary school can have an enjoyable learning experience while learning English with the help of ChatGPT. By talking to ChatGPT in English, they can practice speaking and improve their language skills in a playful way. ChatGPT is like a helpful friend who can also assist with schoolwork. If students have questions about English grammar or need help finding new words, they can ask ChatGPT for guidance. It's a bit like having an interactive AI-based learning companion that loves to help with English homework. Moreover, ChatGPT can create special English lessons just for them (Tlili et al., 2023). It's like having a personalized English teacher who knows exactly what they're interested in. In conclusion, using ChatGPT in English learning can make studying more enjoyable and tailored to each student's interests (Mogavi et al., 2023). However, in the literature, there is not enough evidence specifically regarding the contribution of ChatGPT to the English education of primary school students. This calls for further studies empirically demonstrating the effectiveness of generative AI tools in language learning in primary school settings. Recent research has begun to address this gap, such as a study by Beriak and Mohamad (2026), which demonstrated that integrating ChatGPT voice chat significantly enhances the oral communicative skills of Year 4 ESL pupils. Their findings indicate that AI-driven interaction not only improves grammatical and sociolinguistic competence but also creates a low-pressure environment that reduces speaking anxiety and encourages active participation among young learners. This sort of study is also useful in providing implications for effectively integrating these novel tools into school

settings. Therefore, this study focuses on the impact of ChatGPT on the English reading and writing skills of elementary school students. As a result, answers to the following questions were sought within the scope of the research:

- To what extent does the use of ChatGPT as a digital pen-friend influence the reading skills of primary school English learners?
- How is the impact of ChatGPT-supported instructional activities on the writing proficiency of young EFL learners?
- How does interaction with ChatGPT affect students' attitudes towards English courses compared to traditional teacher-centered methods?

Method

Research Design

A quasi-experimental research design was used in this study. The research design used in the research process is summarized in Table 1. As shown in the table below, the experimental group attended in ChatGPT-supported instructional activities while the control group participated in traditional instructional activities. The *term* “traditional instructional activities” is used to refer to a teacher-centered pedagogical model rather than the specific physical tools (such as chalk or blackboards) used in the classroom. Although the control group made use of modern school infrastructure, including smart boards and digital materials, the overall instructional process followed a conventional pattern in which the teacher functioned as the main source of knowledge and authority. This approach typically involves direct instruction, grammar-based explanations, and a reliance on standardized curriculum materials, which tend to emphasize rote memorization rather than interactive communication (Aydın, 2025). From this perspective, the main difference between the experimental and control groups lies in the mode of pedagogical delivery and the level of student agency—more specifically, the shift from teacher-led instruction to a more personalized, AI-supported form of autonomous learning—rather than simply the presence or absence of digital hardware.

Table 1. Research Design

Groups	Pretest	Manipulation	Posttest
Experimental	Reading Skills Test	ChatGPT-supported instructional activities	Reading Skills Test
	Writing skills Test		Writing skills Test
Control	Attitude Scale	Traditional instructional activities	Attitude Scale

Study Group

The researcher used the convenience sampling technique, a kind of nonprobability or nonrandom sampling. The researchers chose this technique based on the availability of the school where the study was conducted. The participants in the experimental and control groups were non-randomly assigned based on their class membership.

In this study, a total of 59 students attending two separate classes in a primary school constituted the initial study group. Although the school was selected using convenience sampling based on accessibility, the classes themselves were randomly assigned as either the experimental or control group. For the final statistical analysis, four students were excluded: three due to non-participation in pre-tests or post-tests, and one who was identified as an



outlier. This left a final sample of 55 students, with 29 students in the experimental group (13 females, 16 males) and 26 students in the control group (14 females, 12 males).

Instruments

Reading and Writing Skills Test: To separately assess students' reading and writing skills, different types of questions were designed for each skill. For the writing skill, a total of 40 questions were prepared, including fill-in-the-blanks, arranging sentence elements in order, and forming sentences according to given instructions, as stated on the first page of the tests. For the reading skills, a total of 55 questions were prepared, consisting of answering questions based on given passages, fill-in-the-blanks, creating tables, and ordering sentences to form a coherent paragraph. The questions prepared for writing skills were organized as a set of eight questions and the questions prepared for reading skills were organized as a set of six questions. The structure of these questions was designed to appropriately measure students' reading and writing outcomes. Writing questions aimed to evaluate students' grammar and sentence construction abilities, while reading questions were designed to assess comprehension, detail identification, and the ability to make connections within the text. A detailed rubric was used during the scoring process to ensure consistency. The rubric set specific criteria for each type of reading and writing question. Additionally, it was clearly stated which questions were scored for reading and which for writing skills. The reading and writing tests were evaluated separately by two independent evaluators for both the pretest and posttests, and the consistency between the evaluators was examined. Thus, it was ensured that the assessments of reading and writing skills were reliable. The inter-rater consistency coefficients for each question set are demonstrated in Table 2.

Table 2. Inter-rater consistency coefficients

Items	Pre-test Consistency Coefficients	Post-test Consistency Coefficients
Writing 1	0,998	0,998
Writing 2	0,999	0,999
Writing 3	0,999	1,000
Writing 4	1,000	1,000
Writing 5	0,999	0,998
Writing 6	1,000	1,000
Writing 7	1,000	0,999
Writing 8	1,000	1,000
Reading 1	1,000	1,000
Reading 2	1,000	1,000
Reading 3	1,000	0,999
Reading 4	1,000	1,000
Reading 5	0,999	1,000
Reading 6	1,000	1,000

N=55, P<0,001

Table 2 shows that the correlations between the evaluations of the two independent evaluators for both pre-test and post-test scores of all students are between 0.998 and 1.000, which means that the evaluations are largely similar. Accordingly, it was concluded that the evaluation is reliable.

Attitude Scale towards the English course: The Attitude Scale for Elementary English Course developed by Baş (2012) was used to measure whether the use of ChatGPT for educational purposes affects students' attitudes towards the course. When the scale was examined, it was seen that the lower cut-off point of the scale was accepted as 0.32. On the other hand, it was observed that while preparing this scale, there was no restriction on the

number of factors and factors with eigenvalues greater than 1.00 were included in the scale. In the reliability studies of this scale, the eigenvalue was taken as 1.00. The Cronbach Alpha internal consistency coefficient was found to be 0.92. As a result, it was concluded that the scale has a trustworthy structure.

Experimental Process

In this 8-week study, an achievement test and attitude scale were first applied to the experimental and control groups as the pre-tests. Then, every week, two lessons per week, students were given instructions prepared according to the achievements of the annual plan prepared in accordance with the Ministry of Education and Board of Education and Discipline. Students then entered these prompts into ChatGPT and had a dialogue with it about that week's success in a few sentences. Screenshots of the students' dialogues with ChatGPT were regularly submitted to the researcher. At the end of the 8-week study, the achievement tests and attitude scale were administered to the students as the post-test. The pretest and posttests were measured with the rubric prepared by the researcher and the attitude scale.

The lessons were taught the same way in both groups. The resource books and appropriate websites chosen by the teacher were used in the lessons, and the students were given the same homework. Videos on the subject were watched, activities in the book were performed, and mini games were played at the end of the lessons. Both groups were given the same homework assignment to cover that week's learning outcome. The outcomes to be studied for eight weeks are as follows, according to the curriculum:

- Week1: Students will be able to talk about their own and others' possessions.
- Week2: Students will be able to deliver a simple, brief speech about abilities with an initial preparation.
- Week3: Students will be able to understand the general information in an oral text about likes and dislikes.
- Week 4: Students will be able to talk about their likes and dislikes.
- Week 5: Students will be able to engage in simple conversations about likes and dislikes.
- Week 6: Students will be able to ask for clarification in conversations.
- Week 7: Students will be able to talk about their daily routines.
- Week 8: Students will be able to talk about the time.

In-class lessons: Week 1: In the first 10 minutes of the course, students' homework was checked one by one, and then the homework was answered on the smart board. For this week's outcome, "Students will be able to talk about their own and others' possessions", the relevant video was first watched on the school website and then interactive educational games on the website were played. Similarly, in the other weeks, students' homework assignments were checked one by one in the first 10 minutes of the lesson and questions about the homework were answered on the smart board. Then, in accordance with the current curriculum, the topics related to the acquisitions described above were discussed.

ChatGPT activities: The experimental group was given the task of establishing a dialogue with ChatGPT regarding the learning outcome of that week. In the experimental group, instructions were given to ChatGPT in line with the topic to be covered each week, with the aim of students producing dialogues with artificial intelligence. These prompts were given to ensure that ChatGPT gives answers within the scope of the purpose.



Week 1: ChatGPT was given the prompt as follows: "Can you please talk to me like a 9-year-old kid learning English as a second foreign language? Your level must be a beginner. I want to have a dialogue with you using possessive adjectives". In this way, it is aimed to cover the following outcome: "Students will be able to talk about their own and others' possessions", and to chat with the students within the scope of this information (see Figure 1).

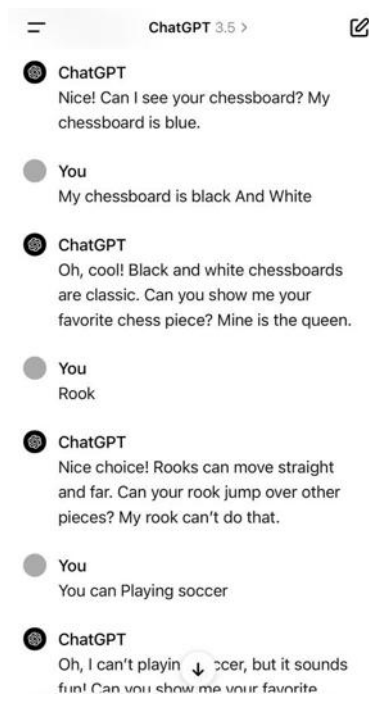


Figure 1. ChatGPT Samples for week 1

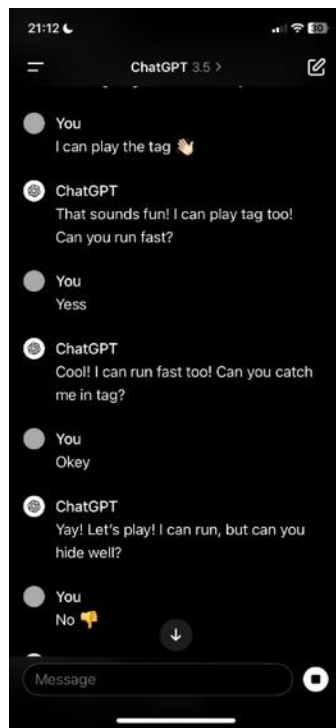


Figure 2. ChatGPT Samples for week 2

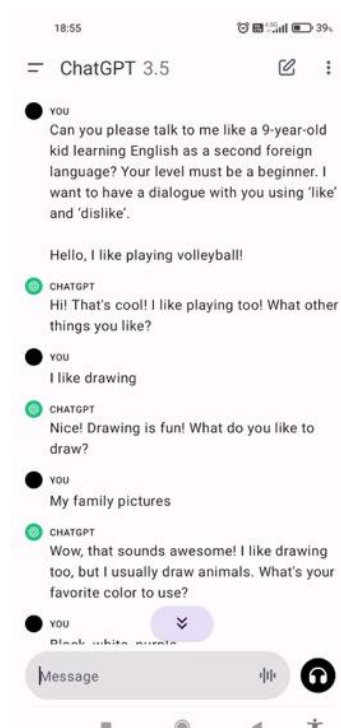


Figure 3. ChatGPT Samples for week 3

Week 2: ChatGPT was given the prompt as follows: "Can you please talk to me like a 9-year-old kid learning English as a second foreign language? Your level must be a beginner. I want to have a dialogue with you using 'can' and 'can't' modals". In this way, it is aimed to cover the "Students will be able to deliver a simple, brief speech about abilities with an initial preparation" outcome and to chat with the students within the scope of this information (see Figure 2).

Week 3: ChatGPT was given the prompt as follows: "Can you please talk to me like a 9-year-old kid learning English as a second foreign language? Your level must be a beginner. I want to have a dialogue with you about free time activities by using 'like' and 'dislike' ". In this way, it is aimed to cover the "Students will be able to understand the general information in an oral text about likes and dislikes." outcome and to chat with the students within the scope of this information (see Figure 3).

Week 4: ChatGPT was given the prompt as follows: "Can you please talk to me like a 9-year-old kid learning English as a second foreign language? Your level must be a beginner. I want to have a dialogue with you using 'like' and 'dislike' ". In this way, it is aimed to cover the "Students will be able to talk about their likes and dislikes." outcome and to chat with the students within the scope of this information (see Figure 4).

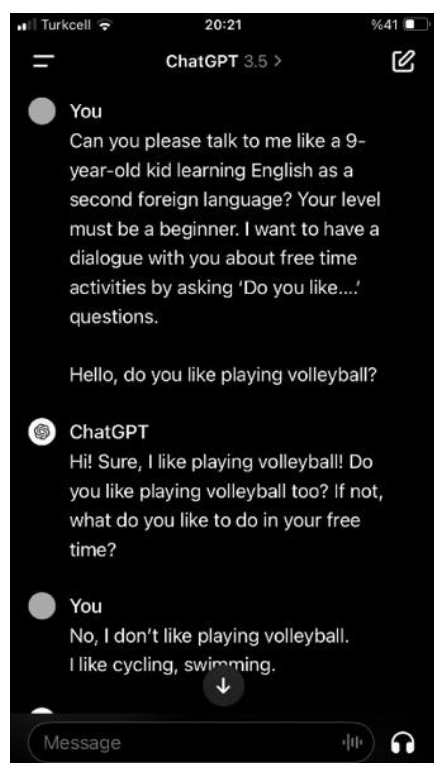


Figure 4. ChatGPT Samples for week 4

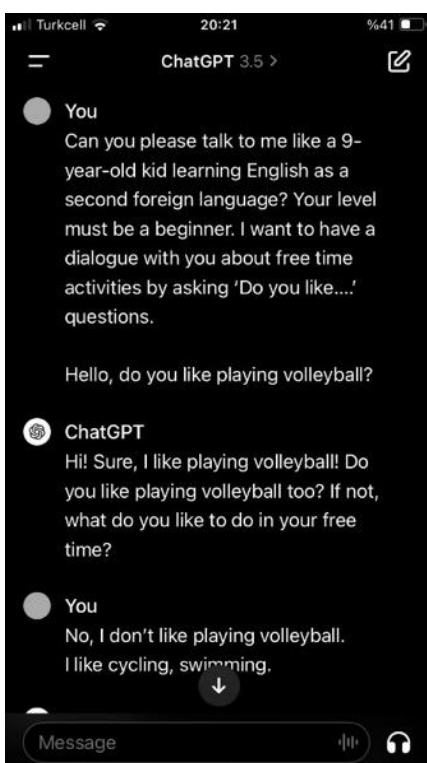


Figure 5. ChatGPT Samples for week 5

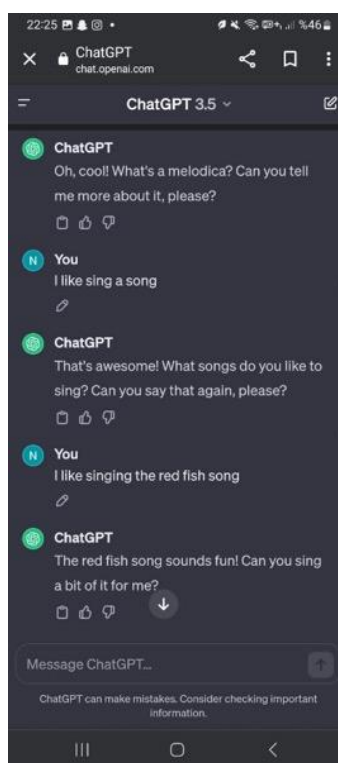


Figure 6. ChatGPT Samples for week 6

Week 5: ChatGPT was given the prompt as follows: "Can you please talk to me like a 9-year-old kid learning English as a second foreign language? Your level must be a beginner. I want to have a dialogue with you about free time activities by asking 'Do you like....' questions". In this way, it is aimed to cover the "Students will be able to engage in simple conversations about likes and dislikes" outcome and to chat with the students within the scope of this information (see Figure 5).

Week 6: ChatGPT was given the prompt as follows: "Can you please talk to me like a 9-year-old kid learning English as a second foreign language? Your level must be a beginner. I want to have a dialogue with you about free time activities by using the clarification sentences like 'Can you say that again, please?', 'Pardon me?', 'Say that again, please'. In this way, it is aimed to cover the "Students will be able to ask for clarification in conversations" outcome and to chat with the students within the scope of this information (see Figure 6).

Week 7: ChatGPT was given the prompt as follows: "Can you please talk to me like a 9-year-old kid learning English as a second foreign language? Your level must be a beginner. I want to have a dialogue with you about basic daily routines". In this way, it is aimed to cover the "Students will be able to talk about their daily routines" outcome and to chat with the students within the scope of this information (see Figure 7).

Week 8: ChatGPT was given the prompt as follows: "Can you please talk to me like a 9-year-old kid learning English as a second foreign language? Your level must be a beginner. I want to have a dialogue with you about our daily routines by making simple inquiries like 'What do you do in the morning?'. In this way, it is aimed to cover the "Students will be able to talk about the time" outcome and to chat with the students within the scope of this information (see Figure 8).



Figure 7. ChatGPT Samples for week 7

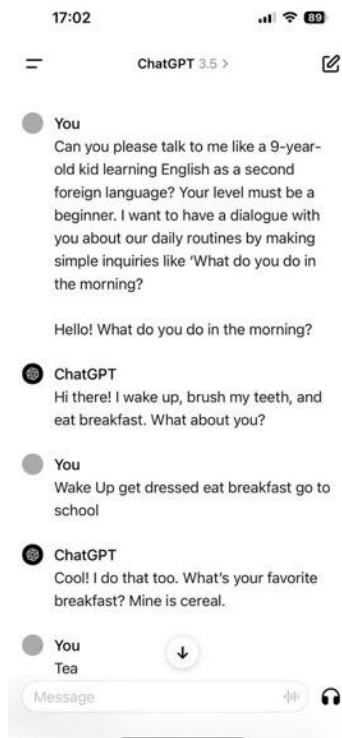


Figure 8. ChatGPT Samples for week 8

Data analyses

In order to determine whether parametric tests can be used in the analysis of the data collected within the scope of the study, the normality of the distribution of the data was assessed using the Shapiro-Wilk Test, and the skewness and kurtosis values were calculated. The results are summarized in Table 3.

Table 3: Findings from the Normality Analyses

		Statistic	df	p	Skewness	Kurtosis
Attitude Total Score	Pre-test	,855	53	,000	-1,343	1,226
	Post-test	,868		,000	-1,495	1,462
Writing Skills	Pre-test	,868		,000	1,067	0,494
	Post-test	,921		,001	0,043	-1,108
Reading Skills	Pre-test	,909		,000	0,427	-0,674
	Post-test	,844		,000	-0,556	-0,346

According to the results of the Shapiro-Wilk Test, the distribution of the scores was found to be non-normal ($p < 0.05$). However, when the skewness and kurtosis coefficients of the scores were examined, these values were found to be between +1.50 and -1.50. According to Büyüköztürk (2008), scales with skewness coefficients between +1.50 and -1.50 can be assumed to follow an approximate normal distribution. The findings regarding the equivalence of the groups before the experimental application are summarized in Table 4.

Table 4. Findings Regarding the Equivalence of the Groups Before the Application

	Group	N	Mean	sd	df	t	p
Attitude Total Score	Cont.	26	108,923	16,555	1,240		0,220
	Exp.	30	102,233	22,771			
Writing Skills	Cont.	26	18,192	14,260	-3,104		0,003
	Exp.	30	35,500	25,125			
Reading Skills	Cont.	26	34,961	18,125	-2,363		0,002
	Exp.	30	49,800	27,197			

Upon examining Table 4, it is observed that the groups were not equivalent in terms of their Writing Skills ($t_{(2-54)}=-3.104$, $p<0.05$) and Reading Skills ($t_{(2-54)}=-2.363$, $p<0.05$) scores before the application with the experimental group scoring significantly higher. To address this initial disparity and ensure a fair comparison of the instructional methods, Multivariate Analysis of Covariance (MANCOVA) was utilized. In this model, pre-test scores for reading, writing, and attitudes were treated as covariates to statistically control for their influence on the post-test results. By adopting this approach, the post-test outcomes were compared using adjusted means, effectively neutralizing the baseline differences and allowing for a more rigorous evaluation of the intervention's impact.

The conditions for meeting the assumptions required for using the MANCOVA test can be summarized as follows:

- There must be two or more dependent variables: Within the scope of the study, three dependent variables were considered—Writing Skills, Reading Skills, and Attitude Toward the English Course.
- There must be a two or more-group independent variable: Within the scope of the study, the independent variable of experimental manipulation consists of two groups—one experimental and one control.
- Observations must be independent: Each student is placed in only one group. Necessary precautions were taken to ensure that participants did not interact with each other while answering the scale.
- The sample size must be adequate: The sample size for each group should be at least 30. However, the literature indicates that this number can be as low as 20 (Büyüköztürk, 2008). The size of both groups within the scope of the study is above 20.
- The reliability of the variable to be controlled must be high: Since there were significant differences between the pre-tests, it is necessary to control for (covary) these pre-tests. In this context, the reliability of the variable to be controlled must be high. The Cronbach's alpha internal consistency coefficient for the attitude scale was determined to be 0.943. For the reading and writing tests, the inter-rater agreement coefficient was considered as the reliability coefficient, and this rate was found to be between 0,998 and 1,000. In this context, it can be said that the reliability of the variable to be controlled is high.
- There should be no univariate or multivariate outliers: Within the scope of the study, outliers in the pre-test and post-test scores of the dependent variables in both the experimental and control groups were investigated using the Mahalanobis Distance test. In this context, the data of one student, who was identified as an outlier, were excluded from the analyses. It was determined that the data met this assumption (Mahalanobis Distance = min: 0.099, max: 8.598).

- Data showing multiple normal distribution: As summarized in Table 2, the collected pre-test and post-test data are normally distributed. In addition, the multivariate normality of the groups formed based on the groups and measurements was also examined, and it was observed that the data followed a normal distribution.
- Multicollinearity: The relationship between the factors and the total scores for both the pre-tests and post-tests was investigated using Pearson's r correlation, and it was determined that there is a significant linear relationship between all factors and the total score, ranging from 0.669 to 0.772 ($p < 0.01$).
- Homogeneity of Regression Slopes: When examining the test of between-subject effects, no significant difference was found between the regression slopes of the variables Group * Writing Pre-test * Reading Pre-test * Attitude Pre-test and the post-test variables (Writing Post-test, Reading Post-test, Attitude Post-test, $p > 0.05$). In other words, when examining the interaction effects between the pre-tests and post-tests, the significance values were found to be greater than 0.05. Accordingly, it was determined that the within-group regression coefficients were equal. In this context, it can be said that the homogeneity of the regression slopes was achieved.
- Homogeneity of Variance-Covariance Matrices: As a result of the Box's test conducted for each factor, the significance level was greater than 0.05 (Box's $M=9.710$, $F=1.518$, $Sig=0.168$), indicating that the assumption of homogeneity of variances was met. Additionally, the Levene test was examined, and it was determined that the variances were equal in terms of the post-test scores.

Findings

When the pre-tests were controlled, the adjusted post-test mean scores for students' English reading and writing skills, as well as their attitudes toward the English course, in the teaching method where ChatGPT was used as a pen friend, are presented in Table 5.

Table 5. Adjusted Post-Test Mean Scores When Pre-Tests Are Controlled

Dependent Variables	Measurement	Pre-test			Post-test		Adjusted Mean	
		n	X	Ss.	X	Ss.	X	SH
Attitude Toward the Course	Experimental	26	108,92	16,55	109,27	16,64	105,64 ^a	2,18
	Control	29	101,28	22,55	106,21	15,53	109,46 ^a	2,05
Writing Skill	Experimental	26	18,19	14,26	44,00	27,68	52,19 ^a	3,80
	Control	29	34,51	24,98	59,14	26,78	51,79 ^a	3,58
Reading Skill	Experimental	26	34,96	18,13	64,85	22,84	70,31 ^a	3,07
	Control	29	49,45	27,61	74,69	18,16	69,79 ^a	2,89

a. Covariates appearing in the model are evaluated at the following values: w.pre = 26,8000, r.pre = 42,6000, attitude pre = 104,8909.

As seen in Table 5, the post-test mean score for writing skills is 52.19 for the experimental group and 51.79 for the control group; the adjusted post-test mean score for reading skills based on the pre-test scores is 70.31 for the experimental group and 69.79 for the control group; the adjusted post-test mean score for attitudes toward the English course based on the pre-test scores is 105.64 for the experimental group and 109.46 for the control group. The MANCOVA results related to the differentiation in post-test scores according to the group covariate when the pre-tests are controlled are summarized in Table 6.

Table 6. The Effect of ChatGPT as a Pen Friend Based on Group

Effect	Wilks' Lambda	F	Hip. sd	Hata sd.	P	η^2	Observed power
Group	,968	0,533	3	48	0,662	0,032	0,151

As shown in Table 6, when the pre-tests are controlled, it was determined that there is no significant difference in the post-test scores between the groups ($F_{(3,48)}=0.533$, $p>0.005$). Accordingly, it can be concluded that the use of ChatGPT as a pen friend in the teaching process did not provide a significantly greater contribution to students' reading and writing skills or their attitudes toward the English course.

Conclusion and Discussion

It has been observed that the use of ChatGPT as a pen friend in the teaching process did not provide a significantly greater contribution to students' reading and writing skills or their attitudes toward the English course. However, both groups showed significant improvement in their reading and writing skills as well as their attitudes after the application, indicating that both groups were successful. One reason for the lack of difference may be that both groups benefited from the process, leading to an overall increase in success levels.

When reviewing the literature and examining studies, studies consistently indicate that ChatGPT has a positive impact on English learners, particularly in enhancing writing skills (Roisah & Widyaningsih, 2024). The incorporation of ChatGPT into educational settings has demonstrated a significant effect on the reading and writing skills of diverse student populations. For example, it has been shown that engineering students' reading comprehension, writing fluency, and general engagement are all improved when ChatGPT is used in conjunction with platforms such as Padlet (Binu, 2024). Similarly, ChatGPT has been shown to enhance English writing abilities at the university level, especially in areas like grammar, coherence, and reasoning, while also encouraging critical thinking (An, 2024). Research has demonstrated that ChatGPT can improve writing quality in high school students as well. Specifically, a pretest-posttest study conducted on Grade 12 students demonstrated that ChatGPT improved content, grammar, and organization (Maghamil & Sieras, 2024).

Although some concerns remain, research on ChatGPT's effect on student attitudes indicates a largely positive reception. For example, Vietnamese university students find ChatGPT to be convenient and helpful, especially when it comes to helping with schoolwork. Although worries about over-reliance and accuracy persist, they appreciate the pedagogical potential of the tool to streamline work and want to utilize it in the future (Le & Tran, 2024). Both their general attitudes toward statistics and their statistical reasoning have improved, according to a study on undergraduate students' attitudes toward using ChatGPT for learning the topic. The application promoted a more assured and involved educational process (Wahba et al., 2024). Students at a technical institution in Ukraine reported improvements in comprehending difficult programming concepts and a modest level of satisfaction with ChatGPT's ability to support their learning. They did, however, admit several drawbacks, such as the need for verification for outdated material, but they nevertheless suggested using it for academic reasons (Chugai & Gavrilenko, 2024). Overall, students in a variety of settings have positive views on ChatGPT because of its educational potential, but they also stress the need for appropriate guidance to minimize its negatives and guarantee its ethical use in the classroom. The majority of the research that is now available focuses on university students, whose learning requirements and capacities are very different from those of younger learners. This approach may help to explain why research on ChatGPT's educational impact has yielded

inconsistent results. Academic challenges at the university level are typically more difficult, enabling deeper feedback interactions that can result in more noticeable advances. Research on EFL learners, for example, shows that AI-assisted feedback improves writing skills greatly by lowering grammatical and lexical errors and organizing text better (Mun, 2024).

Furthermore, the results of the current study can be interpreted through the lens of Cognitive Load Theory. As argued by Alam et al. (2026), while AI-assisted tools have the potential to enhance learning outcomes, they can also impose a significant 'extrinsic cognitive load' on students. For younger learners, such as 4th-grade students, the mental effort required to navigate the AI interface and process its complex linguistic suggestions may create a 'split-attention effect.' This suggests that the cognitive resources of younger learners might be overwhelmed by the dual task of language production and AI interaction, which may explain why difference was observed between the groups in this study. Younger students, however, might need different kinds of help because studies show that although ChatGPT offers insightful feedback, it is not as successful as more conventional approaches for this age group. According to Hutt et al. (2024), older students who could have trouble understanding peer feedback or what makes for high-quality input scored worse on ChatGPT than on classical NLP tools in terms of feedback accuracy. This implies that without pedagogical modifications catered to their developmental stage, integrating ChatGPT for younger learners may not produce the same degree of impact. Conversely, university students typically gain more from AI-supported autonomous learning, which enables more efficient, focused coaching in challenging academic assignments (Ptasznik et al., 2024).

Both ChatGPT users and those employing alternative techniques have shown significant improvements in reading and writing skills, suggesting that various teaching strategies can be effective in the learning process. For example, the Participatory Learning Strategy, which focuses on communicative competencies and emphasizes active learner involvement, has proven effective in enhancing reading and writing skills by transforming students into active participants in the learning process (Tamilselvi et al., 2025). Similarly, a Collaborative and Task-Based Framework, which encourages student autonomy and critical thinking through the integration of authentic, meaningful tasks into literacy activities, has been successful in fostering analytical engagement and significantly improving writing proficiency (Chen et al., 2025). Additionally, Strategic Reading Techniques such as clarifying, predicting, questioning, and summarizing have shown significant improvements in EFL learners' reading skills by providing a structured approach to text comprehension (Okasha, 2020). These alternative techniques further underscore the need for a diverse range of teaching strategies tailored to different age groups and learning environments.

University students typically possess more experience and self-assurance in using digital technologies, which can enhance ChatGPT's effectiveness. By providing individualized and interactive speaking practice, ChatGPT helps students feel more confident. It fills the gaps left by speaking opportunities in traditional classroom settings, thereby improving students' overall language skills (Jingqi & Zong, 2024). While ChatGPT speeds up learning and improves quality, it also makes students feel more comfortable using digital tools. However, in the long run, it may stifle critical thinking and creativity (Md., Awal, & Asaduzzaman, 2024). University students who use ChatGPT report significantly higher levels of academic writing self-efficacy due to the platform's constant feedback, the sense of accomplishment it provides, and the improvement in their composition and language skills (Bouzar, El Idrissi, & Ghourdou, 2024).

In contrast, younger students may not experience the same level of impact. They may also require different types of support regarding assignments and feedback interactions. Consequently, future studies should focus on optimizing the integration of ChatGPT for this age group. Without dispensing with conventional teaching techniques, the chatbot improves younger students' learning experiences and motivation by providing them with constructive criticism and feedback specific to their writing projects (Chang-Yen & Liao, 2023). By providing prompt, personalized responses, encouraging engagement, and enabling individualized learning experiences, ChatGPT can improve the assignment feedback received by younger students, ultimately leading to improved educational outcomes (Zohaib et al., 2024). In conclusion, while ChatGPT holds potential for enhancing reading and writing skills, maximizing this potential will require further research and appropriate strategies tailored to younger learners.

The present study has limitations and recommendations for future research. The first limitation is convenience sampling, which requires further replication studies for the generalizability of the findings from this study. Second, the study was delimited with eight weeks. The length and complexity of the messages students wrote to ChatGPT may have been insufficient to generate the kind of “deep feedback” interaction typically observed in university-level work. This limitation may also be related to the relatively short duration of the intervention. Future studies including longer implementation periods with various instructional methods are required to demonstrate the impact of generative AI on learning outcomes. Besides, longitudinal studies can be employed to track the variation in the impact of generative AI on learning outcomes and other relevant variables. Finally, although experimental studies can display the impact, they cannot reveal learner experience with these novel tools. Thus, future phenomenological studies are highly recommended to reveal the essence of language learners' experiences with ChatGPT and other generative AI tools.

Declarations

Ethical Statement: *The research was conducted in accordance with the relevant ethical rules and academic standards. All procedures performed in studies involving human participants were in accordance with the ethical rules and informed consent was obtained from all individual participants included in the study.*

Funding: *The authors received no financial support for the research and/or authorship of this article.*

Competing Interests: *The authors declare no competing interest. The authors declare that they have meticulously complied with all academic ethical rules during this research.*

Author Contributions: *All authors contributed to the study conception and design. All authors read and approved the final manuscript.*

Data availability and sharing policy: *Data generated or analysed during this study are available from the authors on request.*

References

- Adiguzel, T., Kaya, M. H., & Cansu, F. K. (2023). Revolutionizing education with AI: Exploring the transformative potential of ChatGPT. *Contemporary Educational Technology, 15*(3), ep429, 1-13. <https://doi.org/10.30935/cedtech/13152>
- Aktay, S., Gök, S., & Uzunoğlu, D. (2023). ChatGPT in education. *TAY Journal, 7*(2), 378-406. Retrieved from <https://dergipark.org.tr/en/pub/tayjournal/issue/76658/1259832>
- Alam, M. N., Islam, M. A., Babiker, M. O., Siddiqui, M. S., Amin, M. B., & Oláh, J. (2026). AI-assisted learning tools and student learning outcomes: A cognitive load theory perspective. *Computers in Human Behavior Reports, 21*, 100986.



- An, N. M. (2024). Leveraging ChatGPT for enhancing English writing skills and critical thinking in university freshmen. *Journal of Knowledge Learning and Science Technology*, 3(2), 51-62. <https://doi.org/10.60087/jklst.vol3.n2.p62>
- Awal, M. R. (2024). Curse or blessing? Students' experience from ChatGPT with an application of Colaizzi's phenomenological descriptive method of enquiry and content analysis. *Higher Education, Skills and Work-Based Learning* (ahead-of-print). <https://doi.org/10.1108/HESWBL-09-2023-0249>
- Aydın, S. (2025). Perceived Difficulties among EFL Learners in the Turkish Educational Context. *EJERCONGRESS 2025 Conference Proceedings*, 184.
- Baş, G. (2020). Attitude scale for elementary English course: Validity and reliability study. *International Online Journal of Educational Sciences*, 4(2), 411-424. Retrieved from https://iojes.net/?mod=makale_tr_ozet&makale_id=41211
- Bekou, A., Mhamed, M. B., & Assissou, K. (2024). Exploring opportunities and challenges of using ChatGPT in English language teaching (ELT) in Morocco. *Focus on ELT Journal*, 6(1), 87-106. <https://doi.org/10.14744/felt.6.1.7>
- Beriak, E. I. A., & Mohamad, M. (2026). Exploring Malaysian Upper Primary School ESL Learners' Oral Communicative Skills through ChatGPT Voice Chat. *International Journal of Research and Innovation in Social Science (IJRISS)*, 10(1).
- Bin-Hady, W. R. A., Al-Kadi, A., Hazaea, A., & Ali, J. K. M. (2023). Exploring the dimensions of ChatGPT in English language learning: A global perspective. *Library Hi Tech* (ahead-of-print). <https://doi.org/10.1108/LHT-05-2023-0200>
- Binu (2024). Exploring the efficacy of integrating ChatGPT and Padlet for augmenting reading and writing proficiency among engineering students. <https://doi.org/10.21203/rs.3.rs-4884403/v1>
- Bouzar, A., EL Idrissi, K., & Ghourdou, T. (2024). ChatGPT and academic writing self-efficacy: Unveiling correlations and technological dependency among postgraduate students. *Arab World English Journal (AWEJ) Special Issue on ChatGPT*, April 2024, 225-236. <https://doi.org/10.31235/osf.io/9amk8>
- Büyüköztürk, Ş., Kılıç-Çakmak, E., Akgün, Ö., Karadeniz, Ş., & Demirel, F. (2008). *Bilimsel araştırma yöntemleri [Scientific Research Methods]*. Ankara: Pegem Pub.
- Chang-Yen, L., & Liao, Y. (2023). Design a writing learning companion chatbot based on a knowledge graph to help primary school students practice writing skill. In *Proceedings of 14th IIAI International Congress on Advanced Applied Informatics (IIAI-AAI)* <https://doi.org/10.1109/IIAI-AAI59060.2023.00136>
- Chen, X., Jantharajit, N., & Thongpanit, P. (2025). Enhancing Analytical Reading and Writing Skills in Vocational Education: The Role of Collaborative and Task-Based Learning. *Journal of Education and Learning*, 14(2), 150-158.
- Chen, Z., Wei, W., Zhu, X., & Yao, Y. (2026). Unpacking the rejection of L2 students toward ChatGPT-generated feedback: An explanatory research. *ECNU Review of Education*, 9(1), 20965311241305140.
- Chugai, O., & Havrylenko, K. (2024). ChatGPT: Attitudes and experiences of technical university students in Ukraine. *Information Technologies and Learning Tools*, 101(3), 15-27. <https://doi.org/10.33407/itlt.v10i3.5559>
- Conati, C., & Lallé, S. (2023). Student modeling in open-ended learning environments. In *Handbook of Artificial Intelligence in Education* (pp. 170-183).
- Cronbach, L. J. (1990). *Essentials of psychological testing* (5th ed.). New York: Harper Collins Publishers.
- Dash, B. B., & Gandhi, M. K. (2022). Significance of globalisation and English language. *International Journal on Studies in English Language and Literature (IJSELL)*, 10(5), 10-16. <https://doi.org/10.20431/2347-3134.1005002>

- Ergün, S. (2022). A touch of edtech to create “compelling” blended learning environments. *Journal of Language and Literature Studies in Rumeli*, 31, 1369-1381. <https://doi.org/10.29000/rumelide.1222279>
- Hong, W. C. H. (2023). The impact of ChatGPT on foreign language teaching and learning: opportunities in education and research. *Journal of Educational Technology and Innovation*, 5(1), 37-45. <https://doi.org/10.61414/jeti.v5i1.103>
- Hutt, S., DePiro, A., Wang, J., Rhodes, S., Baker, R. S., Hieb, G., ... & Mills, C. (2024, March). Feedback on feedback: Comparing classic natural language processing and generative ai to evaluate peer feedback. In *Proceedings of the 14th Learning Analytics and Knowledge Conference* (pp. 55-65). <https://doi.org/10.1145/3636555.3636850>
- Javaid, M., Haleem, A., Singh, R. P., Khan, S., & Khan, I. H. (2023). Unlocking the opportunities through ChatGPT Tool towards ameliorating the education system. *BenchCouncil Transactions on Benchmarks, Standards and Evaluations*, 3(2), 100115, 1-12. <https://doi.org/10.1016/j.tbench.2023.100115>
- Jingqi, Z., & Zong, J. (2024). Technology-facilitated oral English learning among Chinese university students: ChatGPT as the digital future. *Journal of Education, Humanities and Social Sciences*, 32, 79-85. <https://doi.org/10.54097/hykh672>
- Kafabih, A., Apriliana, M. S., Putri, N. A., Wardoyo, E., Retnani, S. T., Herda, R. K., & Majid, S. (2026). AI Tools for Personalized English Learning: Bridging the Gap Between Technology and Language Acquisition. *Jurnal Wahana Pendidikan*, 13(1), 1-14.
- Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for language teaching and learning. *RELJ Journal*, 54(2), 537-550. <https://doi.org/10.1177/00336882231162868>
- Kumar, D., Haque, A., Mishra, K., Islam, F., Mishra, B. K., & Ahmad, S. (2023). Exploring the transformative role of artificial intelligence and metaverse in education: A comprehensive review. *Metaverse Basic and Applied Research*, 2, 1-12. <https://doi.org/10.56294/mr202355>
- Le, M. T. T., & Van Tran, K. (2024). Vietnamese university students' perceptions of ChatGPT for homework assistance. *International Journal of Interactive Mobile Technologies*, 18(15), 190-204. <https://doi.org/10.3991/ijim.v18i15.48819>
- Maghamil, M., & Sieras, S. (2024). Impact of ChatGPT on the academic writing quality of senior high school students. *Journal of English Language Teaching and Applied Linguistics*, 6(2), 115-128. <https://doi.org/10.32996/jeltal.2024.6.2.14>
- Menekse, M. (2023). Envisioning the future of learning and teaching engineering in the artificial intelligence era: Opportunities and challenges. *Journal of Engineering Education*, 112(3), 578-582. <https://doi.org/10.1002/jee.20539>
- Mogavi, R. H., Deng, C., Kim, J. J., Zhou, P., Kwon, Y. D., Metwally, A. H. S., ... & Hui, P. (2024). ChatGPT in education: A blessing or a curse? A qualitative study exploring early adopters' utilization and perceptions. *Computers in Human Behavior: Artificial Humans*, 2(1), 100027, 1-20. <https://doi.org/10.1016/j.chbah.2023.100027>
- Mun, C. (2024). EFL learners' English writing feedback and their perception of using ChatGPT. *STEM Journal*, 25(2), 26-45. <https://doi.org/10.16875/stem.2024.25.2.26>
- Okasha, M. A. (2020). Using strategic reading techniques for improving EFL reading skills. *Arab World English Journal*, 11(2), 22-34. <https://doi.org/10.24093/AWEJ/VOL11NO2.22>
- Owan, V. J., Abang, K. B., Idika, D. O., Etta, E. O., & Bassey, B. A. (2023). Exploring the potential of artificial intelligence tools in educational measurement and assessment. *Eurasia Journal of Mathematics, Science and Technology Education*, 19(8), 1-15. <https://doi.org/10.29333/ejmste/13428>

- Ptasznik, B., Wolfer, S., & Lew, R. (2024). A learners' dictionary versus ChatGPT in receptive and productive lexical tasks 37(3), 322-336. *International Journal of Lexicography*. <https://doi.org/10.1093/ijl/ecae011>
- Roisah, S., & Widyaningsih, T. L. (2024). The impact of ChatGPT use on EFL students' writing ability. *Concept: Journal of Social Humanities and Education*, 3(3), 191-199. <https://doi.org/10.55606/concept.v3i3.1403>
- Rugaiyah, R. (2023). The potential of artificial intelligence in improving linguistic competence: A systematic literature review, *Arkus*, 9(2), 319-324. <https://doi.org/10.37275/arkus.v9i2.313>
- Rusmiyanto, R., Huriati, N., Fitriani, N., Tyas, N. K., Rofi'i, A., & Sari, M. N. (2023). The Role of Artificial Intelligence (AI) in developing english language learner's communication skills. *Journal on Education*, 6(1), 750-757. <https://doi.org/10.31004/joe.v6i1.2990>
- Sain, Z. H., Vasudevan, A., Thelma, C. C., & Asfahani, A. (2024). Harnessing ChatGPT for Effective Assessment and Feedback in Education. *Journal of Computer Science and Informatics Engineering*, 3(2), 74-82. <https://doi.org/10.55537/cosie.v3i2.856>
- Tamilselvi, A., Raje, M. S., & Shivani, V. (2025). Participatory Learning—An Effective Strategy to Enhance Communicative Skills. *Journal of Engineering Education Transformations*, 314-321.
- Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments*, 10(1), 15, 1-24. <https://doi.org/10.1186/s40561-023-00237-x>
- Topsakal, O., & Topsakal, E. (2022). Framework for a foreign language teaching software for children utilizing AR, voicebots and ChatGPT (Large Language Models). *The Journal of Cognitive Systems*, 7(2), 33-38. <https://doi.org/10.52876/jcs.1227392>
- Wahba, F., Ajlouni, A. O., & Abumosa, M. A. (2024). The impact of ChatGPT-based learning statistics on undergraduates' statistical reasoning and attitudes toward statistics. *Eurasia Journal of Mathematics, Science and Technology Education*, 20(7), em2468, 1-14. <https://doi.org/10.29333/ejmste/14726>
- 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>