

THE EFFECT OF AI CHATBOTS (E.G CHATGPT) AS A PERSONAL GRAMMAR TUTOR ON GRAMMAR ACHIEVEMENT IN THE WRITING OF ENGLISH EDUCATION STUDENTS

Muhammad Rifky ¹, Juleo Fernando Dwi Surya ², Hanafi ³, Muhammad Ibanez ⁴,
Muhammad Fadlurohman Kaab ⁵, Sujito ⁶

¹⁻⁵ English Language Education Program, 6th Semester

⁶ Department of English Language Education

UIN Raden Mas Said Surakarta

mhmdrifky@gmail.com, juleofermando31@gmail.com, hanafie22@gmail.com,
ibanezr30@gmail.com, kaabagythea66@gmail.com, sujitodoctor@gmail.com

Abstract

This study aimed to examine the effect of ChatGPT as a personal grammar tutor on the grammar achievement of English Education students in narrative writing. The study used a pre-experimental method with a one-group pre-test and post-test design. The participants were 12 undergraduate students from an English Education program. The data were collected through three stages: pre-test, treatment, and post-test. In the pre-test, students wrote a narrative paragraph without AI assistance. During the treatment, students used ChatGPT through structured prompts to identify grammar errors, receive explanations, and learn from corrected sentences. After that, they completed a post-test writing task without using AI tools. The students' writing was assessed using a grammar-focused rubric covering tense consistency, subject-verb agreement, sentence structure, article usage, and punctuation. The results showed that the mean score increased from 10.58 in the pre-test to 14.33 in the post-test, with an average gain of 3.75 points. The paired-sample t-test also showed a significant difference between the scores, $t(11) = 10.68$, $p < .001$. These findings indicate that ChatGPT can support students' grammar development in narrative writing when used as a structured personal grammar tutor.

Article History

Submitted: 16 Juni 2026

Accepted: 19 Juni 2026

Published: 20 Juni 2026

Key Words

ChatGPT, grammar achievement, personal grammar tutor, narrative writing, EFL students.

Introduction

The development of artificial intelligence (AI) technology has significantly affected various aspects of human life, including the educational sector. Within English Language Teaching (ELT), AI-based applications such as adaptive learning systems, intelligent tutoring programs, and conversational chatbots are increasingly utilized to facilitate more individualized learning experiences. These technologies are capable of providing direct feedback related to grammar, pronunciation, and other language competencies, allowing students to learn more independently and efficiently (Crompton & Burke, 2024). Among recent innovations, ChatGPT has gained considerable attention as a conversational AI model that demonstrates strong potential in supporting English language learning across different educational environments (Xiao et al., 2025).

In English as a Foreign Language (EFL) learning, grammatical competence remains a crucial aspect, particularly in academic writing activities. Students' ability to use grammar appropriately strongly influences the clarity, coherence, and accuracy of their written work. Nevertheless, grammar instruction at the university level often encounters several obstacles, including overcrowded classrooms, limited interaction between lecturers and students, and the lack of personalized corrective feedback. These conditions indicate the importance of additional

learning support that students can access beyond classroom instruction. In this regard, AI-based chatbots such as ChatGPT may serve as effective supplementary tools because they can provide grammar explanations, identify language errors, and offer interactive writing practice adjusted to students' individual learning needs. Such features are considered beneficial for improving students' writing performance and grammatical understanding (Song & Song, 2023).

Despite the increasing popularity of AI-assisted language learning, empirical studies examining the role of AI chatbots as personal grammar tutors are still relatively limited. Xiao et al. (2025) reported that previous studies on ChatGPT in EFL writing primarily focused on users' perceptions and attitudes toward the technology, while only a small number investigated its measurable impact on grammar accuracy and language achievement. Similarly, Song and Song (2023) argued that earlier research concerning AI in writing instruction mostly emphasized grammatical error correction rather than broader grammar development. This issue becomes more significant when comparing ChatGPT with applications such as Grammarly, which mainly function as automated correction tools without offering detailed metalinguistic explanations. In contrast, ChatGPT enables more interactive communication through discussions, clarification, and contextual explanations, making it pedagogically richer for language learning purposes (Crompton & Burke, 2024). Furthermore, research conducted by Belda-Medina and Kokošková (2023) explored grammar learning through rule-based chatbot activities; however, the study did not extensively investigate the use of generative AI for natural and dialogue-based grammar tutoring in academic writing settings. This research gap is particularly relevant for English Education students, who are expected not only to master grammar skills for themselves but also to develop competencies as future English teachers. Therefore, further investigation into the influence of ChatGPT as a personal grammar tutor on students' grammar achievement is academically necessary.

Based on the issues discussed above, this study seeks to examine the effect of using ChatGPT as a personal grammar tutor on the grammar achievement of English Education students in writing tasks. The study is grounded in Vygotsky's Sociocultural Theory, which emphasizes the role of a More Knowledgeable Other (MKO) in supporting learners within their Zone of Proximal Development (ZPD). In this context, ChatGPT is positioned as the MKO that assists students during the learning process. A pre-experimental research design involving one group with pre-test and post-test was employed to evaluate students' grammar achievement before and after the treatment. The findings are expected to enrich the existing literature on AI-assisted language learning, provide practical insights for EFL lecturers and curriculum designers, and contribute to future studies on the integration of generative AI technologies in higher education language learning. Ultimately, this research aims to demonstrate that ChatGPT can function not only as a grammar correction application but also as an effective personal tutor that facilitates students' grammatical development in writing.

Method

This research applied a pre-experimental method using a one-group pre-test and post-test design. The design was chosen due to the limited scope of the study, where only one group of participants was involved and observed before as well as after the implementation of the treatment, without the inclusion of a control group. Although this type of design has certain weaknesses related to internal validity, it is still considered suitable for exploratory research that aims to identify initial evidence regarding the effectiveness of an intervention within a manageable and accessible group of participants. In this study, the independent variable refers to the use of

ChatGPT as a personal grammar tutor through structured prompting techniques, whereas the dependent variable is the students' grammar achievement, which is measured through their performance in narrative writing tasks.

The participants of this study consisted of 12 undergraduate students enrolled in the English Education program at a higher education institution. Participants were selected using purposive sampling, based on the criteria that they were actively enrolled in a writing-related course and had prior familiarity with digital tools and internet access. The use of a small participant group is consistent with the pre-experimental scope of this study and is in line with similar small-scale investigations in AI-assisted language learning research. All participants were informed of the study's objectives and their voluntary involvement prior to data collection.

The data collection process was conducted through three consecutive stages: pre-test, treatment, and post-test. These stages were designed to identify the participants' grammar achievement in writing before and after the implementation of ChatGPT as a personal grammar tutor. Therefore, the research procedure was divided into three main stages:

Stage		1				Pre-Test:
-	Narrative	writing	without	AI		assistance
-	1	paragraph	+ -	100	words	theme
-	Theme: My Eid Holiday Experience					
Stage 2 Treatment:						
-	ChatGPT as grammar tutor					
-	Researchers will drop some structured prompts to evaluate pre-test writing.					

Stage 3 Post Test:

- Narrative Writing without AI Assistance
- 1 paragraph + - 100 words
- Theme: A Trip to My Dream City

During the pre-test stage, participants were instructed to write a narrative paragraph independently on the topic "My Eid Holiday Experience" without using AI-based tools, grammar checkers, or spelling applications. The purpose of this activity was to identify the participants' initial grammar competence in a natural writing situation.

In the treatment stage, participants were introduced to several structured prompts designed to use ChatGPT as a personal grammar tutor. The participants then submitted their pre-test writing to ChatGPT and interacted with the AI through guided discussions. In this process, ChatGPT identified grammatical mistakes, explained the related grammar rules, and provided corrected versions of the sentences. Through this interactive feedback process, participants were expected to better understand and apply grammar rules rather than only receiving direct corrections.

Finally, in the post-test stage, participants were required to write another narrative paragraph independently on the topic "A Trip to My Dream City" without any assistance from AI tools. This stage aimed to measure whether there was an improvement in participants' grammar achievement after participating in the ChatGPT tutoring session.

To ensure consistency and pedagogical alignment across all participants, a standardized set of prompts was developed for the treatment session. The prompts were designed to direct ChatGPT to function not merely as an error-correction tool, but as an explanatory grammar tutor that engages the learner in understanding the rationale behind grammatical rules. The following prompt template was used:

Prompt

Act as a personal English grammar tutor. I will give you a paragraph I wrote in English. Please do the following:

1. Identify all grammatical errors in my writing.
2. For each error, explain why it is incorrect using simple language.
3. Provide the corrected version of each sentence.
4. Give me a brief summary of the grammar rules I need to improve.

Here is my paragraph: (Drop Your File !)

The primary instrument used in this study was a writing task administered as both the pre-test and post-test, with different topics to prevent direct memorization or repetition effects. Participants' writing outputs were evaluated using an adapted grammar-focused writing rubric based on the framework proposed by Jacobs et al. (1981), modified to emphasize five grammatical sub-components relevant to narrative text production. The rubric dimensions and scoring criteria are presented in the table below.

Grammar aspect	Description	Score range
Tense consistency	Correct and consistent use of past tense in narrative context	1 (poor) – 4 (excellent)
Subject-verb agreement	Accurate agreement between subject and verb in all clauses	1 – 4
Sentence structure	Variety and correctness of simple, compound, and complex sentences	1 – 4
Article usage	Appropriate use of definite and indefinite articles (a, an, the)	1 – 4
Punctuation	Correct use of commas, full stops, and capitalization	1 – 4

The total grammar score for each participant ranged from 5 to 20 points. Scoring was conducted by the researcher and validated by an inter-rater to ensure reliability. The inter-rater reliability was calculated using Cohen's Kappa coefficient, with a minimum acceptable threshold of 0.70.

The data obtained from the pre-test and post-test grammar scores were analyzed quantitatively. Descriptive statistics, including the mean score, standard deviation, and score gain, were computed to provide an overview of participants' grammar achievement before and after the treatment. To determine whether the difference between pre-test and post-test scores was statistically significant, a paired-sample t-test was conducted using a significance level of $p < 0.05$. Given the small sample size ($n = 12$), the normality of the data distribution was first verified using the Shapiro-Wilk test prior to inferential analysis. All statistical procedures were performed using SPSS software.

Result

This section describes the results of the study based on the grammar scores collected from the pre-test and post-test writing tasks. The participants were 12 undergraduate students from the English Education program. The scores were analyzed quantitatively to see whether using ChatGPT as a personal grammar tutor helped improve the students' grammar achievement in writing narrative texts.

Pre-test Mean	Post-test Mean	Mean Gain	Significance
10.58	14.33	3.75	$p < .001$

Table 1 shows the grammar scores of each participant in the pre-test and post-test. It also includes the score improvement gained by each student after receiving grammar tutoring through ChatGPT.

Participant	Pre-test	Post-test	Gain	%
P-01	11	13	+2	18.2%
P-02	9	12	+3	33.3%
P-03	12	17	+5	41.7%
P-04	10	14	+4	40.0%
P-05	10	13	+3	30.0%
P-06	11	17	+6	54.5%
P-07	9	11	+2	22.2%
P-08	12	16	+4	33.3%
P-09	11	16	+5	45.5%

P-10	10	13	+3	30.0%
P-11	10	14	+4	40.0%
P-12	12	16	+4	33.3%
Mean	10.58	14.33	+3.75	35.4%

Figure 1 compares the pre-test and post-test grammar scores of all participants. The results show that all 12 students improved after the ChatGPT grammar tutoring session. The highest improvement was achieved by P-06, with a gain of 6 points, while the lowest improvement was found in P-01 and P-07, who each gained 2 points.

Table 2 and Figure 2 show the average scores for the five grammar aspects used in the writing rubric. In all aspects, the post-test scores were higher than the pre-test scores. This means that the students showed improvement in different areas of grammar after the treatment. The biggest improvement was found in tense consistency, with a gain of 0.84. This result makes sense because narrative writing mostly requires the use of past tense. Article usage also improved quite well, with a gain of 0.75. Meanwhile, sentence structure had the smallest improvement, with a gain of 0.67.

Before running the paired-sample t-test, the difference scores were first checked for normality using the Shapiro-Wilk test. The result showed that the data were normally distributed, with $W = 0.934$ and $p = .422$. Since the normality assumption was fulfilled, the paired-sample t-test was then used to find out whether the difference between the pre-test and post-test mean scores was significant.

Pre-test mean (SD)	10.58
(1.08)	
Post-test mean (SD)	14.33
(2.02)	
Mean difference	
3.75	
SD of differences	
1.22	
Degrees of freedom	
11	
T-value	
10.68	
Significance (2-tailed)	p
< .001	
Shapiro-Wilk (normality)	W = 0.934,
p = .422	

The paired-sample t-test showed that there was a significant difference between the students' pre-test and post-test grammar scores, $t(11) = 10.68$, $p < .001$. The average post-test score was higher ($M = 14.33$, $SD = 2.02$) than the average pre-test score ($M = 10.58$, $SD = 1.08$), with

an average improvement of 3.75 points. This result suggests that using ChatGPT as a personal grammar tutor helped improve the participants' grammar achievement in narrative writing. All 12 participants showed progress after the treatment, with score gains ranging from 2 to 6 points. Overall, the average percentage of improvement was 35.4%.

Discussion

The results of this study show that using ChatGPT as a personal grammar tutor gave a positive effect on the students' grammar achievement in narrative writing. This can be seen from the increase in the mean score, from 10.58 in the pre-test to 14.33 in the post-test, with a gain of 3.75 points or 35.4%. The paired-sample t-test also showed a significant result, $t(11) = 10.68$, $p < .001$.

This finding indicates that the use of structured interaction with ChatGPT may help students improve their grammar accuracy in writing, especially in a small-scale and short-term study like this. The result is also in line with previous studies about the benefits of AI-assisted tools in EFL writing. For example, Song and Song (2023) found that AI-assisted language learning could improve students' academic writing skills and motivation. Similarly, Crompton and Burke (2024) stated that grammar feedback is one of the useful ways AI can support English language teaching. In this study, ChatGPT was not only used to detect grammar mistakes, but also to explain the errors and give feedback in a more interactive way.

The improvement found in this study can be explained through Vygotsky's Sociocultural Theory (1978), especially the concepts of the Zone of Proximal Development (ZPD) and the More Knowledgeable Other (MKO). In this context, ChatGPT can be seen as a digital MKO because it helped students identify their grammar problems from their pre-test writing. After that, ChatGPT provided explanations, corrections, and examples that supported students in improving their grammar performance within their own learning level. This process is similar to the idea of scaffolding described by Wood, Bruner, and Ross (1976), where learners receive support from someone or something more knowledgeable until they gradually understand the material and become more independent. Through this interaction, students were not only given corrected sentences, but also guidance to understand why the errors happened.

In this study, the interaction with ChatGPT was not only about correcting students' mistakes. The students submitted their own writing and received grammar explanations in simple language, so they could be more involved in understanding the grammar rules. This makes ChatGPT different from automatic grammar checkers such as Grammarly, which mostly only point out errors and give corrections. ChatGPT allowed students to learn through explanation and interaction, not just correction. This is in line with Belda-Medina and Kokošková (2023), who stated that grammar practice through chatbot interaction can help improve learners' accuracy, especially when the chatbot gives feedback that is specific and related to the students' own writing. This ability makes ChatGPT different from earlier rule-based chatbots.

A more detailed look at the grammar rubric scores shows that the improvement was not exactly the same in all five grammar aspects. However, all aspects still showed positive progress after the treatment. These results help explain which grammar areas improved the most and which areas showed smaller improvement after the students used ChatGPT as a grammar tutor.

The biggest improvement was found in tense consistency ($\Delta = +0.84$). This result is quite relevant because narrative writing usually requires students to use the simple past tense consistently. For Indonesian EFL learners, this can be challenging because Bahasa Indonesia does

not have the same tense system as English. In this case, ChatGPT seemed helpful because it could identify tense problems in students' own writing and explain the rules in a way that was easy to understand. Article usage also improved quite well ($\Delta = +0.75$). This shows that ChatGPT's explanation about the use of definite and indefinite articles helped the students, even though articles are often difficult for EFL learners whose first language does not use an article system. On the other hand, sentence structure showed the smallest improvement ($\Delta = +0.67$). This is understandable because writing varied and complex sentences requires deeper understanding and more practice. Therefore, while ChatGPT seems useful for grammar aspects that are more rule-based, more complex writing aspects may need longer and repeated practice to show stronger improvement.

Although all participants showed improvement, the amount of improvement was different for each student. The score gains ranged from +2, which was found in P-01 and P-07, to +6, which was achieved by P-06. This difference shows that ChatGPT may not give the same effect for every learner. Several factors might influence the result, such as students' previous grammar knowledge, how actively they used the prompts, their digital literacy, and the quality of their interaction with ChatGPT during the tutoring session. Students who gained higher scores, such as P-03 and P-06, may have paid more attention to ChatGPT's explanations, asked further questions, or applied the feedback more carefully before doing the post-test. Meanwhile, students with lower gains, such as P-01 and P-07, may have only accepted the corrections without fully understanding the grammar explanations. This finding is related to Zimmerman's Self-Regulated Learning theory (2000), which explains that students' ability to monitor, evaluate, and adjust their own learning strategies can influence their learning outcomes. Therefore, future studies may use think-aloud protocols or interaction logs to better understand how students actually respond to grammar feedback from AI.

The findings of this study also have some implications for EFL teaching, especially in English Education programs in Indonesian universities. First, the results show that ChatGPT can be used as an additional grammar tutoring tool in writing classes. This can be helpful in situations where lecturers have limited time to give individual feedback to every student because of large class sizes or other limitations. However, ChatGPT should not be seen as a replacement for teachers. Instead, it can support the learning process by giving students more chances to get grammar feedback outside the classroom, based on their own pace and needs. Second, since the participants of this study are pre-service English teachers, their experience in using ChatGPT may influence how they view the use of technology in language teaching in the future. After experiencing AI as learners, they may become more open to using similar tools when they become teachers. Third, the structured prompt used in this study can also be adapted by lecturers or curriculum designers. The prompt asks ChatGPT to identify errors, explain the grammar rules, and provide corrected sentences, so it can be used as a practical model for grammar support in undergraduate writing courses.

Although this study provides useful findings, it still has several limitations that need to be considered. These limitations are important to mention so that the results are not interpreted too broadly. They can also become suggestions for future researchers who want to conduct a similar study.

Even with these limitations, this study still gives early evidence that ChatGPT can be used as a personal grammar tutor for EFL university students. For future research, it would be better to involve more participants, use a control group, and apply a mixed-methods design. Researchers can also add qualitative data, such as interviews or analysis of students' interaction logs with

ChatGPT. This can help provide a deeper understanding of how generative AI can support grammar development in EFL learning.

Conclusion

This study focused on how ChatGPT as a personal grammar tutor affected the grammar achievement of English Education students in narrative writing. From the results, the students' grammar scores increased after the treatment. The mean score went up from 10.58 in the pre-test to 14.33 in the post-test, with an average gain of 3.75 points. The paired-sample t-test also showed a significant difference between the pre-test and post-test scores, $t(11) = 10.68$, $p < .001$. This shows that using ChatGPT with structured prompts helped students improve their grammar achievement in writing narrative texts.

Based on the findings, ChatGPT can be considered a helpful additional tool in EFL writing classes. It can help students find their grammar mistakes, understand why the mistakes happen, and learn from the corrected sentences. Even so, ChatGPT should not replace the role of teachers. It is better used as a supporting tool that gives students extra grammar feedback outside the classroom and allows them to study at their own pace.

This study still has some limitations. The number of participants was small, the research only used one group without a control group, and the treatment was conducted in a short time. For that reason, future researchers are suggested to involve more participants, use a control group, and apply a mixed-methods design. Interviews or ChatGPT interaction logs can also be added to see more clearly how students use and respond to AI grammar feedback. Overall, this study shows that ChatGPT has potential to support EFL students in improving their grammar in writing.

Reference

- Belda-Medina, J., & Kokošková, V. (2023). Integrating chatbots in education: Insights from the Chatbot-Human Interaction Satisfaction Model (CHISM). *International Journal of Educational Technology in Higher Education*, 20(1), 1–20. <https://doi.org/10.1186/s41239-023-00432-3>
- Crompton, H., Edmett, A., Ichaporia, N., & Burke, D. (2024). AI and English language teaching: Affordances and challenges. *British Journal of Educational Technology*, 55(6), 2503–2529. <https://doi.org/10.1111/bjet.13460>
- Jacobs, H. L., Zinkgraf, S. A., Wormuth, D. R., Hartfiel, V. F., & Hughey, J. B. (1981). *Testing ESL composition: A practical approach*. Newbury House.
- Song, C., & Song, Y. (2023). Enhancing academic writing skills and motivation: Assessing the efficacy of ChatGPT in AI-assisted language learning for EFL students. *Frontiers in Psychology*, 14, Article 1260843. <https://doi.org/10.3389/fpsyg.2023.1260843>
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.
- Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry*, 17(2), 89–100. <https://doi.org/10.1111/j.1469-7610.1976.tb00381.x>
- Xiao, F., Zhu, S., & Xin, W. (2026). Exploring the landscape of generative AI (ChatGPT)-powered writing instruction in English as a foreign language education: A scoping review. *ECNU Review of Education*, 9(1), 20965311241310881. <https://doi.org/10.1177/20965311241310881>

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13–39). Academic Press.