Effects of Integrating Generative AI Tool on The Writing Skills of Chinese Students toward The Development of an AI-Integrated Teaching Guide

Kurt Pajuelas[1*] and Eirenne Lumasang[2]

[1] National University-Manila, Manila, Philippines
[2] Ateneo de Manila University, Manila, Philippines
kurtpajuelas23@gmail.com

Abstract. This study explores the impact of integrating the generative AI tool ChatGPT on the writing skills of Chinese EFL learners, aiming to lay the groundwork for an AI-integrated teaching guide that upholds essential educational outcomes and fosters autonomous learning. Conducted at the College of Education, Arts, and Sciences (CEAS) at National University - Manila, the study involved 31 Chinese graduate students selected through purposive sampling. Using a mixed-methods approach, a quasi-experimental, one-group pre-test–post-test design measured improvements in grammar, vocabulary, and writing coherence, while semi-structured interviews provided qualitative insights. Quantitative results showed significant gains across all dimensions of writing, including grammar, punctuation, spelling, sentence structure, clarity, consistency, and vocabulary ($p < .001$), confirming the effectiveness of ChatGPT in enhancing productive language skills. Qualitative findings revealed that ChatGPT's real-time, personalized feedback increased learners' language awareness, promoted self-directed learning, and boosted confidence in writing. Synthesizing these results, the study demonstrates that ChatGPT can serve as a powerful complement to traditional instruction, supporting error recognition, skill development, and learner autonomy. The findings inform the design of an AI-driven teaching guide structured around practice, reflection, and independent application to optimize writing instruction in EFL contexts.

Keywords: Educational Technology, EFL Education, Generative AI, Language Acquisition, Writing Proficiency.

1 Introduction

In today's digital age, the integration of technology is reshaping traditional learning methodologies and opening new avenues for learning. With over a billion people studying English as a second or foreign language, language education faces both challenges and opportunities [1]. Artificial Intelligence, now increasingly embedded in workplaces and classrooms, is transforming education [2], [3], [4], [5], with Generative AI tools emerging as significant facilitators of technology-enhanced language learning.

E. Kurniawan et al. (eds.), *Proceedings of the 4th International Conference on English Studies in Indonesia (ICONESIA 2025)*, Advances in Social Science, Education and Humanities Research 1021,
https://doi.org/10.2991/978-2-38476-587-4_8

A range of AI-driven tools support language learners in distinct but complementary ways: Automated Writing Evaluation (AWE) assesses grammar and lexical use; Automated Writing Corrective Feedback (AWCF) offers real-time corrections; AI-enabled Machine Translation (MT) provides quick translations; and Automatic Text Generation (ATG) produces coherent texts [6]. Recent advances have enabled these features to converge in generative AI models like ChatGPT, providing a comprehensive support for language acquisition and writing development.

Central to this study is Schmidt's Noticing Hypothesis [7], [8]. This highlights three key components in language acquisition: awareness, intention, and knowledge. Awareness refers to the learner's ability to notice new linguistic forms—a prerequisite for internalization. AI tools support this by highlighting errors and providing instant feedback. Intention involves the conscious effort to understand and apply language rules. As learners engage with AI tools, they actively correct mistakes and deepen their understanding. Knowledge refers to the accumulation of linguistic information and the ability to apply it correctly. It is built through repeated exposure, allowing learners to reinforce correct usage and minimize future errors.

In the 21st century, AI has become central to educational reform. While some universities limit access to tools like ChatGPT over academic integrity concerns, these tools can support L2 learners, particularly in writing [9], complementing rather than replacing traditional instruction. Successful integration requires planning, training, and ongoing assessment [10]. Generative AI enhances writing through instant, personalized feedback [11], [12], particularly in Asian contexts. In China, the exam-focused, vocabulary-heavy ELT model emphasizes rote memorization over communicative competence, limiting practical language use [13], [14].

Existing research on AI-assisted language learning mainly focuses on high school and undergraduate students, overlooking graduate learners engaged in more discipline-specific writing. Additionally, the pedagogical shift to AI-mediated environments is underexplored, and most studies rely on self-reports rather than performance-based evidence. To address these gaps, this study employs both qualitative and performance-based data, focusing on Chinese graduate students to offer a cross-cultural perspective on AI-assisted EFL learning. It addresses the following questions:

1. Is there a significant improvement in Chinese students' writing skills after being exposed to Generative AI tool ChatGPT?
2. Which features of the generative AI tool significantly influence the students' writing skills?
3. How does noticing influence students' learning autonomy and confidence in writing?
4. What AI-driven teaching guide can be developed based on the findings of the study?

2 Method

The study employs a mixed-methods design, integrating a quasi-experimental one-group pre-test–post-test approach to measure changes in participants' language accuracy. Guided by Schmidt's Noticing Hypothesis [3] [4], it examines how ChatGPT supports language acquisition through functions such as AWE, AWCF, MT, and ATG. Conducted at the College of Education, Arts, and Sciences (CEAS) of National University – Manila, the study involved 31 Chinese graduate students enrolled in the First Term of AY 2024–2025. The study purposively selected participants who had little to no prior interaction with ChatGPT. For the quantitative phase, a Language Accuracy Test adapted from Barrot's assessment was administered as both the pre-test and post-test [15], [16]. In the pre-test, participants wrote a 300+ word argumentative essay on a current social issue within 90 minutes, without reference materials, establishing a baseline in grammar, lexical richness, and writing quality. The six-week intervention phase engaged participants in various language-related tasks with ChatGPT. The post-test mirrored the pre-test in format and conditions but used a different topic to avoid rehearsal effects, enabling rubric-based comparison. Results were analyzed using descriptive statistics. For the qualitative phase, a subset of six participants were selected for semi-structured interviews. The interview protocol included open-ended questions on language error awareness, perceptions of ChatGPT's corrective feedback, and its role in supporting language accuracy. Interview transcripts were analyzed using Thematic Analysis to identify recurring themes [17]. Participants joined voluntarily, with consent procedures ensuring both anonymity and the right to withdraw.

3 Results and Discussion

This study explores the impact of AI-mediated learning on Chinese graduate students studying English as a foreign language (EFL). While previous research has focused primarily on basic education and undergraduate students, this study specifically examines graduate learners. A pre-test–post-test design was employed, in which participants wrote argumentative essays on a particular social issue to assess the effects of AI-assisted writing in practical application of the language.

The results of this study are structured into two main sections: quantitative analysis of the pretest and posttest scores from a one-group, pretest-posttest design with paired samples t-test, and qualitative insights derived from open-ended questions and semi-structured interviews.

3.1 Quantitative Data: Statistical Results and Interpretation

Table 1. Descriptive statistics

	N	Pre-test		Post-test	
		Mean	Standard Deviation	Mean	Standard Deviation
Grammar	31	3.656	0.390	4.344	0.360

	N	Pre-test		Post-test	
		Mean	Standard Deviation	Mean	Standard Deviation
Punctuation	31	3.677	0.417	4.366	0.433
Spelling	31	3.667	0.430	4.312	0.403
Sentence Structure	31	3.667	0.404	4.355	0.421
Clarity	31	3.688	0.471	4.366	0.425
Consistency	31	3.720	0.356	4.344	0.360
Vocabulary	31	3.581	0.421	4.301	0.416
Overall	31	25.656	2.545	30.387	2.521

Table 1 shows notable improvements across all dimensions, indicating a positive impact of the AI tool on students' productive skills. Grammar scores increased from 3.656 (SD = 0.390) to 4.344 (SD = 0.360), while Punctuation rose from 3.677 (SD = 0.417) to 4.366 (SD = 0.433), supporting the notion that ChatGPT aids students in understanding and applying grammatical and punctuation rules through real-time corrections and examples. Spelling improved from 3.667 (SD = 0.430) to 4.312 (SD = 0.403), and Sentence Structure from 3.667 (SD = 0.404) to 4.355 (SD = 0.421), suggesting the benefits of ChatGPT's ability to suggest more complex sentence constructions and provide examples of correct structures. Clarity increased from 3.688 (SD = 0.471) to 4.366 (SD = 0.425). This increase is attributed to ChatGPT's feedback on sentence coherence and language flow. Consistency rose from 3.720 (SD = 0.356) to 4.344 (SD = 0.360), indicating a better grasp of maintaining tense, tone, and personal consistency within writing. Vocabulary improved from 3.581 (SD = 0.421) to 4.301 (SD = 0.416), highlighting exposure to varied lexical choices. Overall, the mean score increased from 25.656 (SD = 2.545) to 30.387 (SD = 2.521), demonstrating the cumulative positive effect of the AI tool across all dimensions of skills.

Table 2. Paired T-test

	Mean	Standard Deviation	Std. Error-Mean	t	df	Sig. (2tailed)	Verbal Interpretation
Grammar	0.688	0.191	0.034	20.039	30.000	0.000	Significant
Punctuation	0.688	0.285	0.051	13.463	30.000	0.000	Significant
Spelling	0.645	0.242	0.044	14.816	30.000	0.000	Significant
Sentence Structure	0.688	0.271	0.049	14.124	30.000	0.000	Significant

	Mean	Standard Deviation	Std. Error-Mean	t	df	Sig. (2tailed)	Verbal Interpretation
Clarity	0.677	0.360	0.065	10.481	30.000	0.000	Significant
Consistency	0.624	0.239	0.043	14.500	30.000	0.000	Significant
Vocabulary	0.720	0.345	0.062	11.616	30.000	0.000	Significant
Overall	4.731	0.823	0.148	32.011	30.000	0.000	Significant

The paired t-test results in Table 2 offer a detailed statistical confirmation of the significant gains in productive language skills among Chinese students after using the Generative AI tool, ChatGPT. Each dimension exhibited notable improvements with p-values below .001, demonstrating statistically significant differences between pre-test and post-test scores across all assessed areas. This strong significance level indicates that the observed gains in language proficiency are directly attributable to the intervention of the AI tools rather than being due to random variation with mean differences highlighting the areas where students improved the most. The cumulative effect of these improvements is reflected in the overall score, which saw an average increase of 4.731 points ($t = 32.011$), highlighting the impact of ChatGPT on language skills.

3.2 Interpretation of Data

Grammar

The pre- to post-test gains suggest that repeated corrective feedback enabled students to internalize grammatical structures [18], [19]. Reductions in errors like subject–verb agreement and tense inconsistency support Noticing Hypothesis, linking repeated exposure to increased awareness and fewer recurrences [7], [8].

Vocabulary

Vocabulary scores improved as ChatGPT's synonym suggestions and alternative phrasing expanded students' lexical range. Research highlights that exposure to varied lexical items and synonyms beyond the formal classroom setting enriches students' vocabulary [20], [21].

Sentence Structure

Initially, participants relied on simple sentences due to limited confidence with complex syntax. ChatGPT's examples and rephrasing suggestions encouraged them to use more sophisticated structures, enhancing the readability and coherence of their writing [22].

Punctuation and Spelling

Students frequently made punctuation errors in the pre-test, but ChatGPT's consistent corrections and explanations led to significant improvements. Repeated exposure increased their attention to proper punctuation, enhancing clarity [23], consistent with research showing that targeted feedback reduces errors over time [24].

Spelling

ChatGPT's real-time corrections reinforced accurate spelling through repeated exposure, enhancing visual recognition and recall. This aligns with findings that AI feedback aids error detection and with cognitive theories linking repeated exposure to long-term retention [25], [26].

3.3 Qualitative Data: Themes from Interview Insights

The qualitative analysis identified several prominent themes from post-intervention interviews, aligned with the Noticing Hypothesis. Thematic analysis revealed insights into how ChatGPT influenced participants' learning experiences that can be seen at Table 3.

Table 3. Statements emerged from the theme

Theme	Statements
Empowerment through Instant Feedback	*"ChatGPT provides clear explanations of grammar rules and real- time feedback, breaking down complex concepts." (Participant 12)* *"I can go over the rules multiple times, and ChatGPT's breakdowns make everything more manageable. I don't feel rushed." (Participant 10)* *"When I'm stuck, I can take as much time as I need without feeling pressured, which helps me understand grammar better." (Participant 7)*
Autonomy in Learning	*"I could ask any question, even if it seemed small or unimportant, and ChatGPT would answer clearly without judgment. This made me more comfortable to explore more challenging grammar rules over time." (Participant 18)* *"It's freeing to be able to ask questions whenever I need. It makes me feel more in control of my learning." (Participant 14)* *"ChatGPT lets me learn at my own pace and curiosity, and I end up learning more because I don't feel any pressure." (Participant 19)*
Enhanced Noticing of Errors	*"I'd review my grammar lessons with ChatGPT after class. Practicing in this way made it easier for me to follow along in class and even answer questions more confidently." (Participant 22)* *"ChatGPT helps me notice mistakes I wouldn't catch on my own, which has made me more aware of my grammar in general." (Participant 13)* *"It's helpful to have a tool that points out my errors and explains them. It makes me more aware of my writing and speaking." (Participant 20)*
Confidence Building Through Error Correction	*"ChatGPT doesn't just say I'm wrong; it shows me why and how to improve. It's like having a patient tutor who believes I can get better." (Participant 5)* *"I usually feel embarrassed making mistakes, but ChatGPT helps me see them as part of learning. I feel like I'm improving with each session." (Participant 16)* *"When I make an error, ChatGPT explains what I missed. This helped me learn to self-correct, which has made me more confident when I write." (Participant 27)*

Theme	Statements
Engagement through Interactive Learning	*"It feels like I'm having a conversation, not just filling in blanks. I can ask ChatGPT about anything I don't understand, and it keeps me interested." (Participant 11)* *"I used to avoid grammar because it felt repetitive. ChatGPT's interactive answers make learning feel fresh every time." (Participant 9)* *"The conversational tone helps me stay focused. I ask questions like I'm talking to a friend, and it makes studying less stressful." (Participant 24)*
Real-world Application through Contextual Examples	*"ChatGPT often gives examples that relate to everyday life, so I can actually see how these grammar rules work in real situations." (Participant 8)* *"I like how ChatGPT explains grammar with sentences I might use in daily conversations. It helps me remember the rules better." (Participant 15)* *"It's easier to learn when I see how grammar rules apply to things I actually say or write. ChatGPT makes grammar feel more useful." (Participant 21)*

Theme 1: Empowerment through Instant Feedback
ChatGPT's real-time, constructive feedback empowers students to notice challenging structures, revisit them until fully understood, and correct errors without feeling discouraged.

Theme 2: Autonomy in Learning
ChatGPT's interactive nature allows students to explore grammar at their own pace, encouraging them to ask questions and direct their attention to areas they find challenging or confusing. This flexibility, combined with the pressure-free learning pace, enables students to deepen their understanding without feeling rushed.

Theme 3: Enhanced Noticing of Errors
ChatGPT enhances students' grammar awareness by not only identifying errors but also explaining the rules behind them, helping learners notice mistakes they might otherwise overlook.

Theme 4: Engagement through Interactive Learning
ChatGPT's conversational style keeps students engaged, allowing them to notice grammar rules through continuous interaction. This repeated exposure reinforces their ability to internalize language structures.

Theme 5: Real-World Application through Contextual Examples
ChatGPT provides everyday language examples, helping students notice grammar rules in real-world contexts. This contextualized approach makes language forms more memorable, relevant, and useful, supporting grammar acquisition through practical examples.

3.4 Cross-Analysis of Quantitative and Qualitative Findings

Students who reported feeling empowered by ChatGPT's feedback were among those who showed the greatest score improvements. Qualitative insights further revealed that participants who focused on specific language areas experienced targeted gains, highlighting the adaptive and personalized nature of AI feedback. These findings also align with Schmidt's Noticing Hypothesis, demonstrating that Generative AI can prompt learners to notice errors, increase awareness, and support active learning.

Table 4. Implementation framework

Phase	AI Role	Role	Assessment Type
PHASE 1: Practice	ChatGPT provides real-time feedback on grammar, vocabulary, and clarity during writing or conversation tasks	Set clear objectives, model appropriate AI use, and monitor student interactions via shared AI links or outputs	Formative review of AI-assisted drafts or conversation transcripts
PHASE 2: Reflection	ChatGPT offers corrections, explanations, and alternative phrasing for revision	Guide students in analyzing and reflecting on ChatGPT's feedback to enhance metacognitive awareness	Formative reflection logs or guided self sessment entries
PHASE 3: Application	AI use limited to planning or brainstorming only (no direct writing assistance)	Evaluate independent writing or speaking tasks, focusing on creativity and critical thinking	Summative pen-and-paper writing or oral tasks

Based on Table 4, an AI-driven teaching guide can be implemented through three progressive phases: (1) Practice, where ChatGPT provides real-time feedback and modeling; (2) Reflection, where students engage with and analyze AI-generated corrections to build awareness and autonomy; and (3) Application, where learners demonstrate unaided performance in creative and critical writing tasks. This framework ensures that ChatGPT supports learning without replacing authentic skill development. The guide should encourage self-paced practice, structured reflection, and teacher facilitation focused on higher-order skills, allowing AI to automate lower-order tasks while fostering self-directed and reflective learning.

4 Conclusion

This study examined the impact of ChatGPT, a generative AI tool, on the writing skills of Chinese graduate students, highlighting its importance in AI-mediated language learning environments. The findings demonstrated that ChatGPT significantly improved students' writing performance across multiple dimensions, including grammar, punctuation, vocabulary, sentence structure, clarity, and consistency. Quantitative data confirmed measurable gains after the six-week intervention, while qualitative insights revealed that ChatGPT's interactive and personalized feedback fostered independent learning, engagement, and confidence. While concerns about overreliance on AI remain in broader educational discussions, the study shows that when integrated thoughtfully, ChatGPT complements traditional instruction and supports meaningful skill development rather than replacing authentic learning. Based on these findings, an AI-driven teaching guide can be implemented in three progressive phases: (1) Practice, where ChatGPT provides real-time feedback on writing and language use; (2) Reflection, where students analyze AI-generated corrections to enhance metacognitive awareness and self-directed learning; and (3) Application, where learners perform tasks independently to demonstrate creativity, critical thinking, and mastery of language skills.

While this study provides strong evidence of the benefits of AI-assisted learning using ChatGPT, several limitations should be acknowledged. First, the sample size was relatively small (n = 31) and limited to Chinese graduate students at a single institution, which may affect the generalizability of the findings. The six-week intervention may not have been long enough to capture long-term retention or sustained writing development. Nonetheless, this study opens several avenues for further research, including examining longer-term interventions to determine whether AI-assisted writing gains are maintained over time and extending investigations to diverse EFL populations across different cultural and institutional contexts.

Acknowledgments. The researchers extend sincere appreciation to Dr. Jessie Barrot for granting permission to adopt the Language Accuracy Test (LAT) he developed, and acknowledge Dr. Elena Mulle, Dr. Jorge Mayordomo, and Dr. Sherlyn Saspa for their expert review, which supported its application in assessing language accuracy improvement among Chinese graduate students.

References

[1]. C. Leacock, M. Chodorow, M. Gamon, and J. Tetreault, *Automated Grammatical Error Detection for Language Learners*, 2nd ed. San Rafael, CA, USA: Morgan & Claypool (Synthesis Lectures on Human Language Technologies, vol. 25), 2014.
[2]. Y. Chen, N. S. Chen, and Y. Lin, "Learning with artificial intelligence: Theories and models," *Educ. Technol. Soc.*, vol. 23, no. 2, pp. 13–24, 2020.
[3]. M. D. Mindzak and S. E. Eaton, "Artificial intelligence in education: A systematic review of literature and mapping of trends," *TechTrends*, vol. 65, no. 4, pp. 526–538, 2021, doi: 10.1007/s11528-021-00611-1.
[4]. N. Naffi, D. Kim, R. Guo, and S. Wang, *Artificial Intelligence in Education: Challenges and Opportunities*. Cham, Switzerland: Springer, 2022.
[5]. S. Srinivasan, "Artificial intelligence in education: The road ahead," *Int. J. Educ. Technol. High. Educ.*, vol. 19, no. 1, Art. no. 3, 2022, doi: 10.1186/s41239-021-00274-5.
[6]. W. Alharbi, "AI in the foreign language classroom: A pedagogical overview of automated writing assistance tools," *Educ. Res. Int.*, pp. 1–15, 2023.
[7]. R. Schmidt, "Attention and awareness in language learning," *Lang. Awareness*, vol. 29, no. 4, pp. 317–330, 2020.
[8]. R. W. Schmidt, "The role of consciousness in second language learning," *Appl. Linguist.*, vol. 11, no. 2, pp. 129–158, 1990, doi: 10.1093/applin/11.2.129.
[9]. J. S. Barrot, "ChatGPT as a language learning tool: An emerging technology report," *Technol. Knowl. Learn.*, vol. 29, pp. 1151–1156, 2023, doi: 10.1007/s10758-023-09711-4.
[10]. D. Aggarwal, "Integration of innovative technological developments and AI with education for an adaptive learning pedagogy," *Catalyst Res.*, vol. 23, no. 2, pp. 709–714, 2023.

[11]. I. D. Basiana, T. Dayrit, S. Estrada, J. M. Lugtu, M. A. Luzung, R. Magalona, and J. C. S. Ana, "Determining the effectiveness of the virtual grammar assistant 'Grammarly' on improving writing skills," *Ignatian Int. J. Multidiscip. Res.*, vol. 2, no. 3, pp. 1210–1223, 2024.

[12]. J. M. F. Carandang, T. L. M. Batica, J. A. T. Echalar, J. A. Pangan, and K. D. Mosende, "An evaluation of online grammar checkers in the development of students' grammar skills: Basis for proposed instructional guidelines," *SDCA Asia-Pac. Multidiscip. Res. J.*, vol. 72, 2023.

[13]. J. Chen, C. H. Lin, G. Chen, and H. Fu, "Individual differences in self-regulated learning profiles of Chinese EFL readers: A sequential explanatory mixed-methods study," *Stud. Second Lang. Acquis.*, vol. 45, no. 4, pp. 955–978, 2023.

[14]. D. Zhang, S. Hennessy, and P. Pérez-Paredes, "An investigation of Chinese EFL learners' acceptance of mobile dictionaries in English language learning," *Comput. Assist. Lang. Learn.*, 2023, doi: 10.1080/09588221.2023.2189915.

[15]. J. S. Barrot, "Trends in automated writing evaluation systems research for teaching, learning, and assessment: A bibliometric analysis," *Educ. Inf. Technol.*, vol. 29, pp. 7155–7159, 2023.

[16]. J. S. Barrot, "Using ChatGPT for second language writing: Pitfalls and potentials," *Assess. Writ.*, vol. 57, Art. no. 100745, 2023.

[17]. V. Braun and V. Clarke, "Using thematic analysis in psychology," *Qual. Res. Psychol.*, vol. 3, no. 2, pp. 77–101, 2006.

[18]. P. Wei, X. Wang, and H. Dong, "The impact of automated writing evaluation on second language writing skills of Chinese EFL learners: A randomized controlled trial," *Front. Psychol.*, vol. 14, 2023, doi: 10.3389/fpsyg.2023.1249991.

[19]. S. Mahapatra, "Impact of ChatGPT on ESL students' academic writing skills: A mixed methods intervention study," *Smart Learn. Environ.*, vol. 11, no. 1, 2024, doi: 10.1186/s40561-024-00295-9.

[20]. A. E. Tantoun, E. A. Elmagasabi, and K. J. Makhluf, "The impact of language exposure on vocabulary acquisition," *مجلة البحوث الأكاديمية (J. Acad. Res.)*, vol. 19, pp. 14–30, 2021.

[21]. A. Al Fraidan and N. Almarri, "Repetition, difficulty, and AI: Enhancing vocabulary acquisition for second language learners," *Afr. J. Biomed. Res.*, vol. 27, pp. 8088–8095, 2024.

[22]. M. Anderson and R. Zhou, "Building complex sentence structures in second language learning through AI feedback," *Lang. Technol. Res.*, vol. 12, no. 4, pp. 214–230, 2020.

[23]. K. Tan and L. Garcia, "Overcoming common punctuation errors in language learning with AI," *Mod. Lang. J.*, vol. 107, no. 2, pp. 289–308, 2023.

[24]. X. Lin and Y. Wei, "Corrective feedback in punctuation use among EFL students," *Lang. Educ. Res.*, vol. 29, no. 3, pp. 317–330, 2020.

[25]. R. M. D. Pratama and D. P. Hastuti, "The use of artificial intelligence to improve EFL students' writing skill," *English Learn. Innov.*, vol. 5, no. 1, pp. 13–25, 2024, doi: 10.22219/englie.v5i1.30212.

[26]. P. Zhang, Y. Sun, and L. Xu, "AI-supported spelling improvement: A longitudinal study," *J. Second Lang. Writ.*, vol. 61, pp. 91–106, 2022.