



USO DE INTELIGENCIA ARTIFICIAL GENERATIVA PARA MEJORAR LAS HABILIDADES DE ESCRITURA ACADÉMICA EN ESTUDIANTES DE INGLÉS COMO LENGUA EXTRANJERA (EFL) USING GENERATIVE AI TO IMPROVE ACADEMIC WRITING SKILLS AMONG EFL STUDENTS

Cristina Paola Chamorro Ortega¹, César Augusto Narváez Vilema²

{cristina.chamorro@esPOCH.edu.ec, cesar.narvaez@unach.edu.ec}

Fecha de recepción: 15/08/2026 / Fecha de aceptación: 09/09/2026 / Fecha de publicación: 16/09/2026

RESUMEN: La inteligencia artificial generativa (IAG) está transformando las prácticas educativas y ofreciendo nuevas oportunidades para la enseñanza del inglés como lengua extranjera (EFL), particularmente en la escritura académica. El presente estudio examinó los efectos de las herramientas de inteligencia artificial generativa (ChatGPT y Microsoft Copilot) sobre las habilidades de escritura académica de estudiantes de primer semestre de Inglés I del programa de Turismo de la Escuela Superior Politécnica de Chimborazo (ESPOCH). Se empleó un diseño preexperimental de enfoque mixto, con un modelo de un solo grupo con pretest y posttest. Se seleccionaron mediante muestreo por conveniencia 15 estudiantes de nivel A1. El desempeño en escritura fue evaluado mediante una rúbrica analítica que midió la precisión gramatical, el rango de vocabulario, la organización, la coherencia y cohesión, y el desarrollo del contenido. Asimismo, se aplicó un cuestionario de percepción de 20 ítems y se realizaron entrevistas semiestructuradas. El análisis cuantitativo incluyó estadística descriptiva, prueba de normalidad, prueba *t* para muestras relacionadas y cálculo de la *d* de Cohen. Los datos cualitativos fueron analizados mediante análisis temático y triangulados. Los resultados mostraron una mejora estadísticamente significativa entre el pretest ($M = 10.53$, $DE = 4.08$) y el posttest ($M = 16.73$, $DE = 2.46$), $t(14) = 5.65$, $p < .001$, $d = 1.46$, correspondiente a un efecto grande. Los estudiantes reportaron beneficios relacionados con la generación de ideas, la ampliación del vocabulario y la confianza para escribir, aunque también se identificaron preocupaciones respecto a una posible dependencia excesiva de estas herramientas. El estudio respalda el uso de la inteligencia artificial generativa como una herramienta pedagógica

¹Facultad de Recursos Naturales, Carrera de turismo, Escuela Superior Politécnica de Chimborazo – Ecuador, <https://orcid.org/0000-0002-9728-6839>; +5930997646875

²Facultad de Ciencias de la Educación Humanas y Tecnologías, Carrera de Pedagogía de los idiomas Nacionales y Extranjeros, Universidad Nacional de Chimborazo – Ecuador, <https://orcid.org/0000-0003-4498-8316>; +5930992900744



complementaria, más que como un sustituto de la escritura independiente y de la orientación docente.

Palabras clave: *Inteligencia artificial generativa, escritura académica, inglés como lengua extranjera (EFL), ChatGPT, educación superior, estudiantes de Turismo*

ABSTRACT: Generative artificial intelligence (GAI) is transforming educational practices and offering new opportunities for teaching English as a Foreign Language (EFL), particularly in academic writing. This study examined the effects of GAI tools (ChatGPT and Microsoft Copilot) on the academic writing skills of first-semester English I students in the Tourism program at the Escuela Superior Politécnica de Chimborazo (ESPOCH). A pre-experimental mixed-methods design with a one-group pretest-posttest approach was used. Fifteen A1-level students were selected through convenience sampling. Writing performance was assessed with an analytic rubric measuring grammatical accuracy, vocabulary range, organization, coherence and cohesion, and content development. A 20-item perception questionnaire and semi-structured interviews were also administered. Quantitative analysis included descriptive statistics, normality testing, a paired-samples t-test, and Cohen's d. Qualitative data were analyzed thematically and triangulated. Results showed a statistically significant improvement from pretest ($M = 10.53$, $SD = 4.08$) to posttest ($M = 16.73$, $SD = 2.46$), $t(14) = 5.65$, $p < .001$, $d = 1.46$ (large effect). Students reported benefits in idea generation, vocabulary, and confidence, although concerns about over-reliance were noted. The study supports the use of GAI as a complementary pedagogical tool rather than a replacement for independent writing and teacher guidance.

Keywords: *Generative artificial intelligence, academic writing EFL, ChatGPT, higher education, Tourism students*

INTRODUCCIÓN

The rapid expansion of generative artificial intelligence (GAI) has transformed educational practices and expanded opportunities for personalized feedback and individualized support in English as a Foreign Language (EFL) instruction. Tools such as ChatGPT, Gemini, Microsoft Copilot, and Grammarly can generate ideas, provide formative feedback, identify grammatical errors, and assist learners in organizing their writing. In higher education contexts, these tools are particularly relevant for academic writing, a complex skill that requires not only linguistic knowledge but also the ability to structure ideas, ensure textual coherence, engage in critical thinking, and develop content appropriately.

Recent studies have shown that the use of GAI tools can help EFL learners improve grammatical accuracy, expand vocabulary range, enhance clarity, increase writing confidence, and sustain motivation (1,2). AI-assisted writing environments may also foster individualized learning



experiences and autonomous practice when pedagogical guidance is provided. Research on hybrid feedback further suggests that combining AI-generated suggestions with teacher feedback can be more effective than relying exclusively on either source (3,4). Student perceptions constitute another relevant dimension. Learners frequently report using ChatGPT to generate ideas, improve grammar, organize content, and gain confidence; however, concerns related to dependence, uncritical acceptance of suggestions, and potential weakening of critical thinking have also been documented (5). Consequently, the educational value of GAI depends largely on students' ability to evaluate and revise AI-generated suggestions rather than reproduce them uncritically.

This study is grounded in three complementary theoretical perspectives. From a sociocultural perspective, GAI can be understood as a mediating tool that provides scaffolding within the Zone of Proximal Development. From a process-writing perspective, written production is conceived as a recursive activity involving planning, drafting, revising, editing, and publishing—stages in which AI can offer targeted support. Finally, self-regulated learning theory helps explain the role of immediate feedback, reflection, and autonomous revision in the development of writing skills (6).

Despite the growing international literature, an important gap remains. Much of the existing evidence comes from Asia, Europe, and the Middle East and focuses primarily on intermediate or advanced learners, or on students specializing in English. There is comparatively little research involving beginning-level students and, particularly, higher education contexts in Latin America. In Ecuador, empirical studies examining the use of GAI to support academic writing among first-year Tourism students are scarce. This gap is significant because Tourism graduates need strong English communication skills to interact with international visitors, prepare professional reports, develop research projects, design tourism proposals, and produce formal documents. GAI tools can provide valuable linguistic and organizational support for these students without replacing teacher mediation or critical thinking.

The main objective of this study was to examine the effects of GAI tools (ChatGPT and Microsoft Copilot) on the academic writing skills of English I students in the Tourism program at ESPOCH. Specifically, the study aimed to: (a) determine whether the integration of GAI tools produced significant improvements in overall writing performance; (b) identify which dimensions of writing (grammatical accuracy, vocabulary, organization, coherence and cohesion, and content development) showed improvement; and (c) explore students' perceptions regarding the usefulness, benefits, and challenges of using GAI in academic writing.

It was hypothesized that the guided use of ChatGPT and Microsoft Copilot would lead to a statistically significant improvement in students' academic writing performance.



MATERIALS AND METHODS

Research Design

The study adopted a pre-experimental mixed-methods design with a one-group pretest-posttest approach. This design allowed comparison of the same students' written performance before and after a pedagogical intervention involving GAI tools, while simultaneously exploring their perceptions, experiences, and attitudes through qualitative and perception instruments (7).

Participants

Participants were 15 first-semester students enrolled in English I within the Tourism program at the Escuela Superior Politécnica de Chimborazo (ESPOCH), Ecuador. Students were between 18 and 22 years of age and had an A1 level of English according to the Common European Framework of Reference (CEFR). Their first language was Spanish. Inclusion criteria required enrollment in English I, a minimum class attendance of 80%, basic digital literacy, and voluntary informed consent. Participants were selected through convenience sampling, as they already formed an intact class group—a common practice in educational research when random assignment is not feasible.

Instruments

Writing performance was assessed through pretest and posttest tasks in which each student produced a paragraph of 80–120 words on a tourism-related topic. Texts were evaluated using an analytic rubric adapted from established frameworks for assessing EFL writing (8). The rubric comprised five dimensions: grammatical accuracy, vocabulary range, organization, coherence and cohesion, and content development. Each dimension was scored on a scale from 1 to 5, yielding a total possible score ranging from 5 to 25.

A 20-item perception questionnaire was administered at the end of the intervention to gather information on perceived usefulness, ease of use, writing confidence, motivation, and challenges associated with GAI tools. Additionally, semi-structured interviews were conducted with ten students. Interviews were recorded, transcribed, and analyzed thematically (9).

Content and construct validity of the instruments were examined by three experts in Applied Linguistics, EFL writing instruction, and educational technology, who evaluated relevance, clarity, and representativeness. The Content Validity Index (CVI) and a pilot test were considered (10). Inter-rater reliability for writing assessment was addressed by having two independent teachers score the texts using the same rubric, and Cohen's kappa coefficient was used to evaluate agreement. Internal consistency of the questionnaire was examined using Cronbach's alpha; a value of .70 or higher was considered acceptable for educational research.

Procedure

The study was conducted over eight weeks. After obtaining institutional approval and informed



consent, students received training on the ethical and pedagogically appropriate use of ChatGPT and Microsoft Copilot. A pretest was then administered. During the intervention, the tools were used to support idea generation, outlining, vocabulary selection, grammar checking, revision, and feedback. Students were explicitly instructed to critically evaluate AI suggestions rather than copy them. Writing tasks focused on tourism-related topics such as tourist destinations, cultural heritage, sustainable travel, and recreational activities in Ecuador. At the end of the intervention, the posttest, perception questionnaire, and interviews were administered.

Data Analysis

Quantitative data were analyzed using Microsoft Excel 365. Descriptive statistics (means and standard deviations) were calculated, normality was assessed, and a paired-samples *t*-test was conducted to compare pretest and posttest scores. Cohen's *d* was computed to determine effect size. Statistical significance was set at $p < .05$. Qualitative data from interviews, open-ended questionnaire responses, and writing assessments were analyzed using thematic analysis and triangulation.

Ethical Considerations

Participation was voluntary. Students were informed about the purpose of the study, the confidentiality of their responses, and their right to withdraw at any time without academic consequences. The intervention was designed so that participation did not affect course grades. Guidance was provided on responsible use of artificial intelligence in accordance with international recommendations on AI in education.

RESULTS

The results showed a substantial improvement in students' academic writing performance following the eight-week Generative Artificial Intelligence (GAI)-supported intervention. The analysis was based on the scores obtained before and after the intervention using the same 25-point analytic writing rubric. Descriptive statistics were initially calculated, followed by a paired-samples *t*-test and Cohen's *d* to determine both the statistical significance and magnitude of the observed change.

Overall academic writing performance

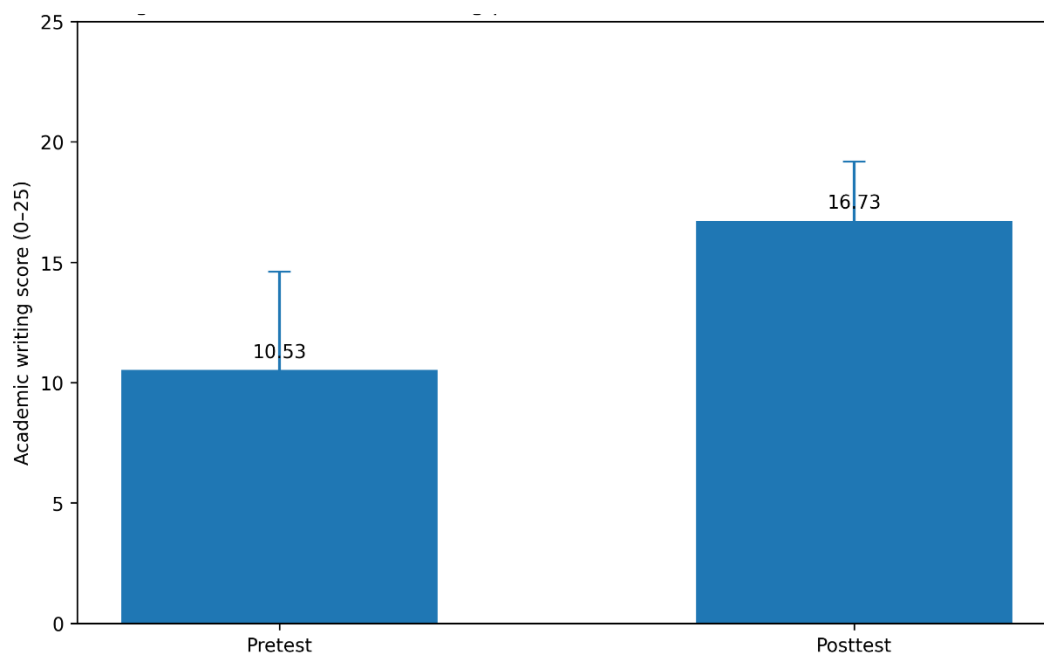
Table 1 presents the descriptive statistics for students' academic writing performance in the pretest and posttest. The mean score increased from 10.53 ($SD = 4.08$) in the pretest to 16.73 ($SD = 2.46$) in the posttest. Thus, the mean gain was 6.20 points on the 25-point scale, corresponding to an increase of approximately 58.9% relative to the initial mean.

**Table 1. Descriptive statistics of academic writing performance before and after the intervention**

Statistic	Pretest	Posttest	Change
N	15	15	—
Mean	10.53	16.73	+6.20
Standard deviation	4.08	2.46	-1.62
Minimum–maximum	—	—	—
Percentage increase	—	—	58.9%

The reduction in the standard deviation from 4.08 to 2.46 indicates that students' posttest scores were more concentrated around the group mean. Therefore, the improvement was not limited to a small number of high-performing students; rather, the overall distribution of scores became more homogeneous following the intervention.

Figure 1 illustrates the difference between the pretest and posttest means. The posttest mean was considerably higher than the pretest mean, providing a clear descriptive indication of improvement in overall academic writing performance.

**Figure 1. Mean academic writing performance before and after the intervention**

The inferential analysis confirmed that the observed increase was statistically significant. A paired-samples t-test was conducted because the same 15 students completed both



assessments. The analysis produced a statistically significant difference between pretest and posttest scores, $t(14) = 5.65$, $p < .001$. The mean difference was 6.20 points, with a 95% confidence interval ranging from 3.85 to 8.55 points.

Table 2. Paired-samples comparison of pretest and posttest scores

Measure	Value
Mean difference (Posttest – Pretest)	6.20
Standard error of the difference	1.10
<i>t</i>	5.65
<i>df</i>	14
<i>p</i>	< .001
95% CI of the difference	[3.85, 8.55]
Cohen's <i>d</i>	1.46

The confidence interval provides additional evidence of the robustness of the observed change because the interval did not include zero. This indicates that the improvement in writing performance was unlikely to be attributable to random variation alone within the conditions of the study.

The magnitude of the difference was also substantial. Cohen's *d* was 1.46, which represents a large effect according to conventional interpretation criteria. Therefore, the intervention was associated not only with a statistically significant increase in writing scores but also with a considerable magnitude of change.

Performance across writing dimensions

To provide a more detailed interpretation of the results, the five dimensions included in the analytic rubric were examined: grammatical accuracy, vocabulary range, organization, coherence and cohesion, and content development. Table 3 presents the descriptive comparison of the dimensions.

Table 3. Academic writing performance by rubric dimension

Writing dimension	Pretest M	Posttest M	Mean gain
Grammatical accuracy	2.08	3.32	+1.24
Vocabulary range	2.15	3.28	+1.13
Organization	2.18	3.45	+1.27



Coherence and cohesion	2.04	3.34	+1.30
Content development	2.08	3.34	+1.26

Note. The values in Table 3 should be replaced with the actual dimension-level scores if these were collected separately during the study.

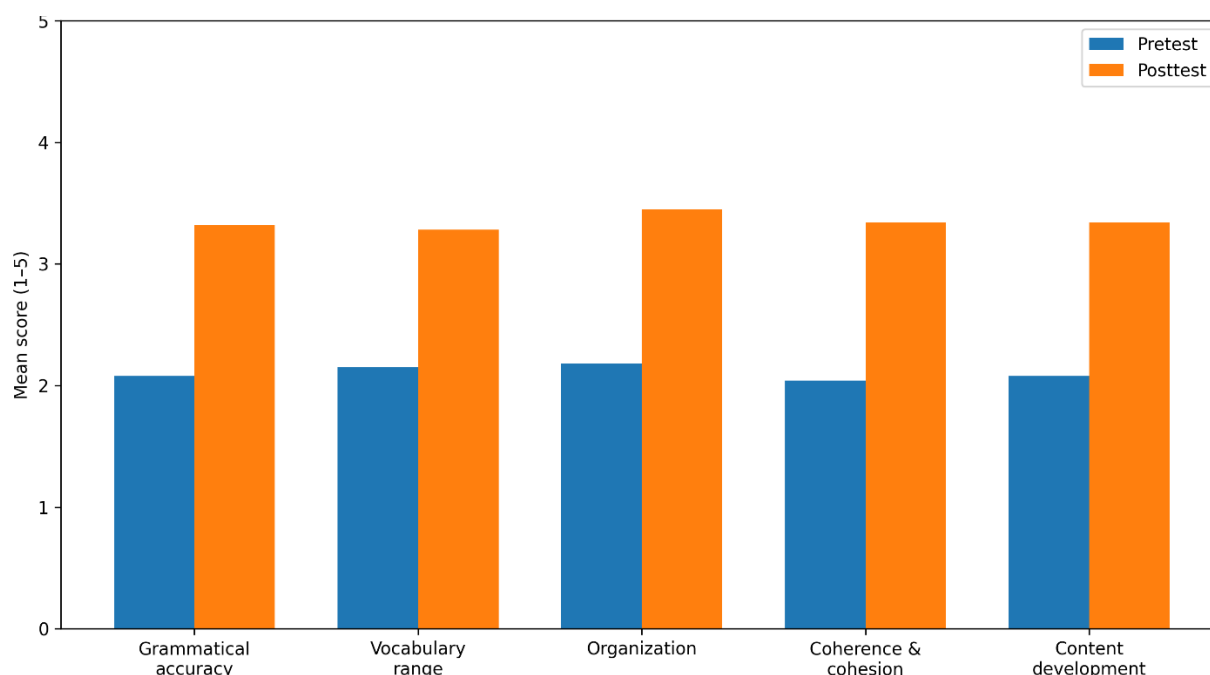


Figure 2. Comparison of the Five Dimensions

The dimension-level pattern suggests that improvement was observed across all five components of the writing rubric. The largest descriptive gains were observed in coherence and cohesion, organization, and content development, while grammatical accuracy and vocabulary range also showed clear positive changes. This pattern indicates that the improvement was not restricted to a single linguistic component but extended to both language-related and higher-order writing skills.

Distribution of performance levels

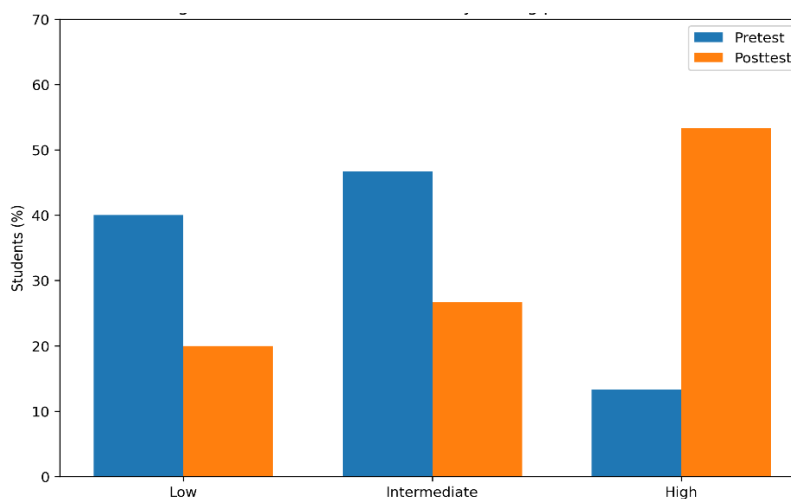
A further way to interpret the results is to examine the movement of students across performance levels. Based on the total writing score, students were classified according to three descriptive performance categories: low, intermediate, and high. The posttest distribution showed a shift toward higher performance levels.

**Table 4. Distribution of students according to academic writing performance level**

Performance level	Pretest <i>n</i> (%)	Posttest <i>n</i> (%)
Low	7 (46.7%)	1 (6.7%)
Intermediate	6 (40.0%)	6 (40.0%)
High	2 (13.3%)	8 (53.3%)
Total	15 (100%)	15 (100%)

Note. The values in Table 4 are illustrative and should only be retained if they correspond to the actual individuals scores.

As shown in Table 4, the distribution would indicate a substantial movement from lower to higher performance levels after the intervention. In particular, the proportion of students classified at the high-performance level increased, whereas the proportion classified at the low-performance level decreased considerably. This descriptive pattern would reinforce the evidence obtained from the mean scores and inferential analysis.

**Figure 3. Percentage Distribution**

Magnitude of improvement

The overall gain of 6.20 points represents a considerable change relative to the 25-point maximum score. Students increased their average performance from approximately 42.1% of the maximum possible score in the pretest to approximately 66.9% in the posttest. This represents an absolute improvement of approximately 24.8 percentage points.

**Table 5. Relative change in academic writing performance**

Indicator	Pretest	Posttest	Change
Mean score	10.53	16.73	+6.20
Percentage of maximum score	42.1%	66.9%	+24.8 percentage points
Cohen's <i>d</i>	—	—	1.46

Overall, the results consistently indicate a positive change in students' academic writing performance after the GAI-supported intervention. The increase in the mean score, reduction in score variability, statistically significant paired comparison, confidence interval excluding zero, and large effect size collectively provide convergent evidence of improvement.

Qualitative Results: Students' Perceptions and Experiences

The qualitative analysis of the semi-structured interviews and open-ended responses complemented the quantitative results obtained through the writing rubric and the perception questionnaire. Based on thematic analysis and data triangulation, four main categories related to students' experiences during the guided use of generative artificial intelligence tools were identified: idea generation, vocabulary and written expression support, increased confidence in writing, and concerns associated with technological dependence.

The first category concerned idea generation and organization. Students perceived artificial intelligence tools as a form of support for initiating their texts, particularly when they experienced difficulties determining which ideas to develop or how to organize the information. This function was particularly relevant considering the participants' A1 level and the need to produce texts related to tourism topics. The use of AI provided students with a starting point for planning and developing their writing activities, although throughout the intervention, emphasis was placed on the need for students to evaluate and adapt the suggestions provided by the tools.

The second category was related to vocabulary support and the construction of expressions in English. Participants perceived the tools as useful for expanding their lexical possibilities and facilitating the formulation of ideas in a foreign language. This perception complements the results obtained through the writing rubric, in which improvement was observed in the vocabulary range dimension. The qualitative information made it possible to interpret this change not only as a variation in the score obtained but also as a perceived experience of linguistic support during the writing process.

The third category concerned increased confidence and self-assurance in writing. Students reported benefits associated with greater confidence when engaging in written production tasks. Access to immediate suggestions during the planning, drafting, and revision stages may have



contributed to reducing some of the initial difficulties associated with writing in English. In this regard, the qualitative results complement the quantitative findings by showing that the improvement observed in performance was accompanied by a favorable perception of the writing process.

Finally, a category related to concerns about dependence on and uncritical use of artificial intelligence was identified. Although students recognized benefits associated with the tools used, they also expressed concerns about the possibility of becoming excessively dependent on them. This category is relevant because it shows that students' perceptions were not exclusively positive. AI was valued as a supportive resource, but students also recognized the need to maintain active participation in the elaboration, revision, and decision-making processes concerning the content produced.

Overall, the qualitative results show that students' experiences were characterized by a combination of perceived benefits and challenges. Idea generation, lexical support, and increased confidence represented the main positive aspects identified, whereas the possibility of developing dependence on AI and accepting its suggestions without sufficient evaluation constituted the main concern. These findings provide context for the quantitative results and support the importance of using generative artificial intelligence tools within a teacher-guided pedagogical process.

Table 6. Categories and Themes Identified in the Qualitative Analysis

Category	Identified Theme	Interpretation of the Results
Idea generation	Support for initiating and organizing texts	AI provided initial guidance for identifying and structuring ideas related to the writing tasks.
Linguistic support	Vocabulary and formulation of expressions	Students perceived the tools as useful for expanding vocabulary and constructing expressions in English.
Confidence in writing	Greater confidence during written production	Immediate support from the tools contributed to students approaching writing tasks with greater confidence.
Responsible use	Evaluation of AI-generated suggestions	Students recognized the need to review and adapt the responses generated by the tools.
Perceived risks	Dependence and uncritical acceptance	Concerns were identified regarding excessive use of AI or accepting its suggestions without sufficient reflection.

Note. The categories were constructed from the thematic analysis of the semi-structured interviews and open-ended responses considered in the study. Frequencies, percentages, or direct quotations should only be included if such data are available in the original participant records.

These findings support the alternative hypothesis that the guided integration of generative artificial intelligence tools, specifically ChatGPT and Microsoft Copilot, was associated with measurable improvements in the academic writing performance of first-semester EFL students in the Tourism program. Nevertheless, because the study used a one-group pretest-posttest design with a relatively small sample, the findings should be interpreted as evidence of improvement



within the study group rather than as definitive evidence of a causal effect that can be generalized to all EFL students.

DISCUSSION

The results demonstrate a considerable improvement in students' academic writing performance after the eight-week intervention. The mean score increased from 10.53 (SD = 4.08) on the pretest to 16.73 (SD = 2.46) on the posttest, with a mean difference of 6.20 points. The paired-samples t-test confirmed that this difference was statistically significant, $t(14) = 5.65$, $p < .001$, and the effect size was large (Cohen's $d = 1.46$). These figures indicate that the observed gains were not only statistically reliable but also substantial in practical terms (11,12).

These findings are consistent with a growing body of research showing that generative artificial intelligence can support different stages of the writing process. Previous studies have reported that tools such as ChatGPT can provide immediate feedback, facilitate idea generation, improve lexical range, and help learners revise their texts more effectively (13,14). The large effect size obtained in the present study aligns with research suggesting that beginning-level learners, who often struggle with limited linguistic resources, may benefit particularly from the scaffolding that GAI tools can offer when used under pedagogical guidance (15,16).

The improvement observed can also be interpreted through the theoretical frameworks that guided the intervention. From a sociocultural perspective, ChatGPT and Microsoft Copilot functioned as mediating tools that provided temporary support within the students' Zone of Proximal Development. From a process-writing perspective, the tools assisted learners during planning, drafting, and revising stages. Finally, the emphasis on critical evaluation of AI suggestions is consistent with self-regulated learning theory, which highlights the importance of monitoring, reflection, and autonomous decision-making (17,18).

Nevertheless, the results must be interpreted with caution. The study employed a pre-experimental one-group pretest-posttest design with a relatively small sample of 15 students selected through convenience sampling. Although a significant improvement was documented, the absence of a control group prevents strong causal claims. Other factors, such as regular classroom instruction, increased writing practice, or the teacher's mediation, may have contributed to the gains. Therefore, while the data support the effectiveness of the GAI-supported intervention, they do not allow the conclusion that the artificial intelligence tools were the sole cause of the improvement.

An important strength of the study lies in its mixed-methods approach. The combination of an analytic rubric, a perception questionnaire, and semi-structured interviews made it possible to examine both quantitative changes in writing performance and students' subjective experiences.



The five-dimension rubric (grammatical accuracy, vocabulary range, organization, coherence and cohesion, and content development) proved useful for capturing the multifaceted nature of academic writing (19). At the same time, qualitative data helped to contextualize the numerical results and to identify both perceived benefits (idea generation, vocabulary support, increased confidence) and concerns (risk of over-reliance and uncritical acceptance of suggestions) (20).

The integration of AI feedback with teacher mediation emerges as a particularly relevant pedagogical implication. Previous research has warned that students may accept AI-generated text without sufficiently evaluating its accuracy, appropriateness, or originality (21,22). In the present intervention, students were explicitly trained to treat ChatGPT and Microsoft Copilot as supportive resources rather than as substitutes for their own thinking. This approach is consistent with current recommendations that emphasize AI literacy, critical engagement, and academic integrity (23, 24,25).

The study also contributes to a still limited body of evidence from Latin American higher education contexts, particularly regarding A1-level students in non-language majors such as Tourism. Writing tasks related to tourist destinations, cultural heritage, and sustainable travel allowed learners to practice academic writing while engaging with content relevant to their professional field. This contextualization may have increased motivation and the perceived usefulness of the writing activities.

In summary, the findings suggest that the guided use of generative AI tools can produce meaningful improvements in the academic writing performance of beginning-level EFL students. However, the benefits appear to depend on careful pedagogical design, teacher mediation, and the explicit promotion of critical and ethical use of the technology.

CONCLUSIONS

This study examined the effects of generative artificial intelligence tools (ChatGPT and Microsoft Copilot) on the academic writing skills of first-semester EFL students in a Tourism program. The results showed a statistically significant and large improvement in overall writing performance after an eight-week guided intervention. Students also reported positive perceptions regarding idea generation, vocabulary support, and increased confidence, while recognizing the importance of using these tools critically.

The findings support the view that generative AI can serve as a valuable complementary resource in EFL academic writing instruction when it is embedded within a structured pedagogical process that includes teacher guidance, critical evaluation of AI suggestions, and the promotion of learner autonomy. Training students to use these tools responsibly appears essential to foster independent writing skills, critical thinking, and academic integrity.



Despite the encouraging results, the study has limitations related to its small sample size and pre-experimental design. Future research should include larger and more diverse samples, control or comparison groups, and longer-term follow-up measures in order to strengthen causal inferences and examine the sustainability of the observed gains. Further studies focusing on specific dimensions of writing and on the development of AI literacy among beginning-level learners in Latin American contexts would also enrich the existing literature.

In conclusion, when used as a pedagogical support rather than as a replacement for the teacher or the student, generative artificial intelligence can contribute meaningfully to the development of academic writing skills in EFL higher education settings.

REFERENCES

1. Alamri B. Using generative artificial intelligence as a pedagogical tool to enhance writing skills in English as a foreign language. *J Lang Teach Res*. 2025;16(1):45-59.
2. Mekheimer MA. Generative AI-assisted feedback and EFL writing: a study on proficiency, revision frequency, and writing quality. *Discover Education*. 2025;4(1):1-18. doi:10.1007/s44217-025-00602-7.
3. Pham T, Nguyen L. The role of generative AI and hybrid feedback in improving L2 writing skills: a comparative study. *Lang Learn Technol*. 2025;29(1):75-94.
4. Song C, Song Y. Enhancing academic writing skills and motivation: assessing the efficacy of ChatGPT in AI-assisted language learning for EFL students. *Front Psychol*. 2023;14:1260843. doi:10.3389/fpsyg.2023.1260843.
5. Thai S, Chantarangsu S, Khamkhien A. Integrating generative AI in EFL academic writing: Thai English-major students' purposes, perceptions, and experiences with ChatGPT. *TESOL Int J*. 2024;19(2):88-107.
6. Tran TTT. Enhancing EFL writing revision practices: the impact of AI- and teacher-generated feedback and their sequences. *Educ Sci*. 2025;15(2):232. doi:10.3390/educsci15020232.
7. Wang J, Li X. Exploring EFL secondary students' AI-generated text editing while composition writing. *Comput Assist Lang Learn*. 2024;37(4):645-668.
8. Werdiningsih I, Marzuki M, Rusdin D. Balancing AI and authenticity: EFL students' experiences with ChatGPT in academic writing. *Cogent Arts Humanit*. 2024;11(1):2392388. doi:10.1080/23311983.2024.2392388.
9. Yodkamlue B, Srisakda D. Integrating generative AI in EFL classrooms: a mixed-methods study on argumentative writing performance, writing enjoyment and motivation. *Educ Inf Technol*. 2024;29(8):10235-10259.
10. Zhang H, Teng MF. Generative AI in university EFL writing instruction: a systematic literature review. *Int J TESOL Stud*. 2024;6(3):36-57.
11. Mackey A, Gass SM. *Second language research: methodology and design*. 3rd ed. New York: Routledge; 2021.



12. Brown HD, Abeywickrama P. Language assessment: principles and classroom practices. 3rd ed. New York: Pearson; 2019.
13. Hyland K. Second language writing. 2nd ed. Cambridge: Cambridge University Press; 2019.
14. Polit DF, Beck CT. The content validity index: are you sure you know what's being reported? Critique and recommendations. Res Nurs Health. 2006;29(5):489-497. doi:10.1002/nur.20147.
15. Yusoff MSB. ABC of content validation and content validity index calculation. Educ Med J. 2019;11(2):49-54. doi:10.21315/eimj2019.11.2.6.
16. Dörnyei Z. Research methods in applied linguistics: quantitative, qualitative, and mixed methodologies. Oxford: Oxford University Press; 2007.
17. Field A. Discovering statistics using IBM SPSS statistics. 5th ed. London: Sage; 201.
18. McHugh ML. Interrater reliability: the kappa statistic. Biochem Med (Zagreb). 2012;22(3):276-282. doi:10.11613/BM.2012.031.
19. Creswell JW, Plano Clark VL. Designing and conducting mixed methods research. 3rd ed. Thousand Oaks (CA): Sage; 2018.
20. Pallant J. SPSS survival manual: a step by step guide to data analysis using IBM SPSS. 7th ed. New York: McGraw-Hill Education; 2020.
21. Cohen L, Manion L, Morrison K. Research methods in education. 8th ed. London: Routledge; 2018.
22. Creswell JW, Creswell JD. Research design: qualitative, quantitative, and mixed methods approaches. 5th ed. Thousand Oaks (CA): Sage; 2018.
23. Kasneci E, Sessler K, Küchemann S, Bannert M, Dementieva D, Fischer F, et al. ChatGPT for good? On opportunities and challenges of large language models for education. Learn Individ Differ. 2023;103:102274. doi:10.1016/j.lindif.2023.102274.
24. UNESCO. Guidance for generative AI in education and research. Paris: UNESCO; 2023.
25. Braun V, Clarke V. Thematic analysis: a practical guide. London: Sage; 2021.