

Integrating ChatGPT into University English Teaching: Linguistic Performance, Autonomy, and Learning Experience

Jeany Argueta¹, Gisela Espinoza² and Víctor Sandoval²

¹Investigación, Universidad Gerardo Barrios, San Miguel, El Salvador

²Desarrollo e Innovación, Universidad Gerardo Barrios, San Miguel, El Salvador

jlis22@ugb.edu.sv

ggarcia@ugb.edu.sv

victorsandoval@ugb.edu.sv

Abstract: This study examines how the planned integration of ChatGPT influences English language development, student autonomy, and the overall learning experience in a university English as a Foreign Language (EFL) Intermediate English II course in El Salvador. The aim is to analyze both quantifiable changes in language proficiency and students' interactions with artificial intelligence as a learning support tool, particularly in relation to self-regulation and independent learning processes. Furthermore, the study sought to understand how students perceive the usefulness and limitations of AI in academic contexts. A convergent mixed-methods case study design was implemented with twenty-three undergraduate students enrolled in an Intermediate English II course. Quantitative data were collected through TOEFL ITP simulation tests administered before and after the intervention to assess changes in language proficiency. Qualitative data were collected through a focus group of four participants, where students shared their perceptions, attitudes, and experiences interacting with ChatGPT during the learning process. Both datasets were analyzed separately and then integrated to provide a comprehensive interpretation of the findings. The quantitative results reveal a positive, though not statistically significant, trend in TOEFL ITP scores after the intervention. While quantifiable improvements in test performance were limited, the qualitative findings offer a clearer understanding of the learning process. Students reported using ChatGPT as an accessible and ongoing learning partner to practice language skills, clarify doubts, generate examples, and receive immediate feedback. This consistent interaction contributed to increased confidence, greater autonomy, and the development of self-regulated learning strategies. However, participants also expressed concerns about potential over-reliance on AI and occasional inaccuracies in the generated responses, underscoring the importance of maintaining critical awareness. This standalone study highlights the pedagogical potential of AI tools when intentionally integrated into language teaching and underscores the need to foster critical AI literacy so that such technologies enhance, rather than hinder, meaningful learning.

Keywords: TOEFL ITP, Artificial Intelligence, Higher Education, Learner Autonomy, Second Language Acquisition

1. Introduction

There are many debates on the rapid development of generative artificial intelligence (AI) and the digital transformation of higher education. Dempere et al. (2023) identify key benefits of ChatGPT in higher education, such as research assistance, automated grading, and improved human computer interaction, as well as concerns regarding academic integrity, online testing security, job displacement, digital literacy inequalities, and AI induced anxiety.

The integration of AI tools such as ChatGPT raises important theoretical considerations related to learner autonomy. Conceptualized by Holec (1981) as the learner's ability to take charge of decisions concerning their learning. Autonomy has since been expanded by (Little, 1991; Benson, 2011) to include socially mediated and technologically supported dimensions

Recent research shows the potential of generative artificial intelligence tools to support language learning through feedback, personalized learning, practice activities, and facilitate interactive use, potentially enhancing students' engagement and independent learning (Kasneci *et al.*, 2023; Kohnke, Moorhouse and Zou, 2023) In the context of English as a Foreign Language (EFL), emerging research suggests that AI may support vocabulary development, writing practice, and grammar awareness by providing on-demand explanations and personalized language input (Yan, 2025). However, empirical evidence regarding the impact of these tools on broader language proficiency and learner autonomy remains limited. In addition, much of the existing research has been conducted in Asian and European educational contexts, leaving a gap in empirical studies examining the pedagogical integration of generative AI in Latin American universities and its relationship to measurable language development and student learning experiences.

TOEFL reports show that El Salvador ranked 47th globally in 2025 English proficiency (Education First, 2025), highlighting systemic challenges in English acquisition. At the institutional level, Universidad Gerardo Barrios reported a 3.6% failure rate in intermediate and advanced English courses, suggesting that a portion of students

continue to experience difficulties in achieving expected linguistic competencies. These factors point to the need for innovative yet pedagogically grounded interventions.

This study investigates the linguistic and experiential dimensions of ChatGPT with the following questions:

- What changes are observed in the group's linguistic performance throughout the semester?
- How do students perceive the usefulness, limitations, and academic ethics of using ChatGPT?
- What evidence of autonomy appears, and how is it related to the use of ChatGPT?

2. Background

The integration of Artificial Intelligence (AI) into higher education has generated growing interest in its potential to enhance language learning outcomes. In the context of English as a Foreign Language (EFL), technological tools are increasingly being used to support skill development, provide personalized feedback, and promote student autonomy.

2.1 English as a Foreign Language (EFL) in Higher Education

English has become the dominant language of academic communication, scientific publishing, and professional interaction worldwide (Arquez Muñoz, Tobón, and Cely Betancourt, 2025). In higher education institutions, English proficiency is increasingly positioned as a prerequisite for meeting institutional standards related to teaching quality, research productivity, and internationalization, particularly in non-English speaking contexts where mobility, access to academic literature, publication expectations, and accreditation processes often operate through English (Naka and Spahija, 2022).

Learning English as a Foreign Language (EFL) presents specific challenges, particularly in contexts where exposure to authentic language input is limited (Atrianus Toni, 2026). Within higher education, this global dominance increases the demands placed on students' academic writing and speaking, as EFL learners often face not only linguistic gaps but also cognitive and strategic difficulties, as well as affective barriers such as anxiety, factors that can limit performance even when learners understand course content (Pu and Chang, 2025; Wijaya, 2026). Empirical evidence from TOEFL ITP contexts further demonstrates that students' listening difficulties are influenced by both external factors, such as speaker accent, speed of delivery, and complex sentence structures, and internal factors, including limited vocabulary, lack of practice, concentration problems, and test-related anxiety, all of which contribute to persistently low proficiency scores in standardized assessments (Nuraida, 2025).

Consequently, although English proficiency is considered essential for academic success and institutional development, many students of English as a foreign language in higher education continue to face structural, linguistic, and affective constraints that prevent them from fully participating in academic practices.

2.2 TOEFL ITP as a Standardized Measure of English Proficiency

The TOEFL ITP (Institutional Testing Program), developed by the Educational Testing Service (ETS), is a widely used standardized assessment designed to measure English proficiency in academic contexts (ETS, 2026). The test evaluates three core components: Listening Comprehension, Structure and Written Expression, and Reading Comprehension. Scores range from 310 to 677 and provide a general indicator of language proficiency.

Standardized assessments such as TOEFL are designed to ensure reliability and validity (Khan, Imran, and Javed, 2025). Reliability refers to the consistency of scores, while validity refers to the degree to which the test measures academic English proficiency. Given its recognized reliability, validity, and predictive power, the TOEFL ITP is a standardized tool for assessing academic English proficiency and for conducting pre- and post-test assessments in the field of educational research.

2.3 Artificial Intelligence in Education

Artificial intelligence (AI) refers to a field of computer science focused on creating systems that can perform tasks typically associated with human intelligence, such as learning from experience, reasoning, solving problems, perceiving information, and processing language (Allen, 2020; Long and Magerko, 2020; Stryker and Kavlakoglu, 2025).

Adaptive learning systems adjust the difficulty of content based on student needs, providing immediate feedback and modifying instructional pace according to individual performance (Yang, 2025). Recent research also indicates that, while AI helps improve personalized learning and student engagement, its use must be carefully regulated to avoid overdependence and potential negative effects on critical thinking and academic integrity (Vieriu and Petrea, 2025).

When thoughtfully implemented, AI technologies can enhance learner autonomy and support differentiated instruction, while still requiring human oversight to ensure pedagogical and ethical alignment (Holmes, Bialik and Fadel, 2019). Therefore, the use of IA tools in education should not be understood as a substitute for teaching practices, but rather as a complementary tool to improve teaching and learning.

2.4 Artificial Intelligence in English Language Learning

AI applications in language learning are commonly applied in areas such as automated grammar and writing, speech recognition systems that improve pronunciation, intonation, and fluency, and conversational agents that provide simulated communicative interactions (Haciyeva, 2025). In EFL contexts, these technologies increase exposure to the English language and offer individualized practice opportunities that are often limited in traditional classrooms.

Artificial intelligence offers multiple benefits in foreign language learning, notably its ability to provide immediate feedback, personalize learning, offer continuous practice opportunities, provide adaptive learning experiences, and improve language skills such as grammar, vocabulary, and speaking (AlTwijri and Alghizzi, 2024; Bounif, 2026). They typically provide targeted instruction, automated feedback, and interactive activities that help students practice English more frequently and in a variety of contexts.

While AI technologies offer promising tools to support language development, their effectiveness will largely depend on how they are incorporated into teaching practices and how educators guide them. Therefore, AI should be considered a complementary resource for improving English language learning, rather than a substitute for teacher-led instruction and human interaction in the classroom.

2.5 Learner Autonomy

Learner autonomy was originally conceptualized by Holec, (1981) as “the ability to take charge of one’s own learning”, a definition that has become foundational in language education research. This concept implies that learners are responsible for every key aspect of the learning process, such as setting goals, selecting methods and strategies, and monitoring progress and evaluating outcomes (Little, 1991).

From this perspective, autonomy is not an innate trait, but a capacity that must be developed with pedagogical support. Holec emphasized that the shift of responsibility from the teacher to the learner represents a fundamental transformation in educational practices, positioning learners as active agents in their own learning process rather than passive recipients of knowledge.

In the context of language learning, student autonomy is related to self-directed learning, metacognitive awareness, and reflective practices. Autonomous students demonstrate the ability to make decisions about their own learning, adapt strategies to their needs, and engage in ongoing self-assessment. Consequently, the development of learner autonomy has been linked to greater motivation, more effective learning outcomes, and the promotion of lifelong learning skills.

3. Methodology

The study employed a convergent mixed methods case study designed to examine the integration of ChatGPT in a university Intermediate English II course. The focus was on a single cohort of students enrolled at Gerardo Barrios University in San Miguel, El Salvador.

3.1 Quantitative

The quantitative segment of this research follows a non-experimental, longitudinal, prospective design, involving two measurements throughout the process. The study adopts a relational scope, aimed at comparing the results obtained in the initial phase (pre-test) with those from the final stage (post-test). To achieve this, the quantitative data analysis was structured into three phases. First, an exploratory analysis was conducted, involving the descriptive treatment of the data and graphical representation of the variables (Hernández-

Sampieri and Mendoza, 2018). Second, a correlational analysis was performed to compare input and output results, based on the assumption of a positive trend in final performance (Okoye and Hosseini, 2024). Third, a hypothesis test was carried out using inferential statistics for related samples to determine the statistical significance of the observed differences (Al-kassab, 2022).

3.2 Qualitative

The qualitative element of this study explored students' perceptions of the integration of ChatGPT during their Intermediate English B2 Course. A focus group of four participants was conducted, and the data were analysed using thematic analysis following the approach proposed by Braun and Clarke (2006). The semi-structured interview held at the end of the semester allowed participants to reflect on their language learning experiences, the ways in which the tool supported their autonomy, and ethical considerations related to the use of AI.

The analysis followed these six phases: (1) Familiarization with data; (2) Initial coding; (3) Searching for themes; (4) Reviewing themes; (5) Defining themes; and (6) Producing report (Braun and Clarke, 2006).

Following this procedure, the coding process revealed several categories that were subsequently grouped into broader themes related to students' prior familiarity with AI tools, learner autonomy, and perceived limitations and risks of AI use. The creation of a codebook registered codes, definitions, excerpts, and themes ensures transparency and consistency during the analytical process.

4. Data Analysis

This section presents the findings organized according to the research questions, integrating quantitative and qualitative data. Tables and figures are included to support the interpretation of results and provide a clearer visualization of the findings.

4.1 Changes in Students' Linguistic Performance

RQ1: What changes are observed in the group's linguistic performance throughout the semester?

The quantitative analysis shows a general upward trend in students' linguistic performance across all components of the TOEFL ITP test. Mean scores increased from entry to exit levels in Listening Comprehension (41.4 to 42.8), Structure and Written Expression (39.4 to 41.0), Reading Comprehension (40.5 to 40.7), and the overall TOEFL ITP Level (404.0 to 415.1). These results suggest a tendency toward improvement in students' English proficiency over the course of the semester.

Table 1: Comparative table of descriptive statistics (entry vs. exit)

Statistical Parameter	Listening Comprehension		Structure and Written Expression		Reading Comprehension		TOEFL ITP Level	
	Entry	Exit	Entry	Exit	Entry	Exit	Entry	Exit
mean	41.4	42.8	39.4	41.0	40.5	40.7	404.0	415.1
std	7.4	6.7	6.0	5.9	5.9	5.3	54.7	53.2
min	31.0	33.0	32.0	33.0	32.0	33.0	333.0	340.0
25%	35.5	38.0	35.0	37.0	35.5	36.5	367.0	378.3
50%	40.0	42.8	38.0	41.0	39.0	40.7	397.0	415.1
75%	45.0	46.5	41.5	42.5	45.0	44.5	423.5	428.3
max	62.0	65.0	57.0	61.0	56.0	55.0	583.0	603.3

Note: There is a general upward trend in the mean scores across all components from the Entry Level to the Exit Level. This suggests a positive impact of the intervention or the period between the two assessments.

To further examine the distribution of scores, Figure 2 compares entry and exit levels. The findings suggest that although mean scores increased, the variability in performance also changed. This indicates that not all students improved at the same rate, and the distribution of scores at the exit level differs slightly from that at entry.

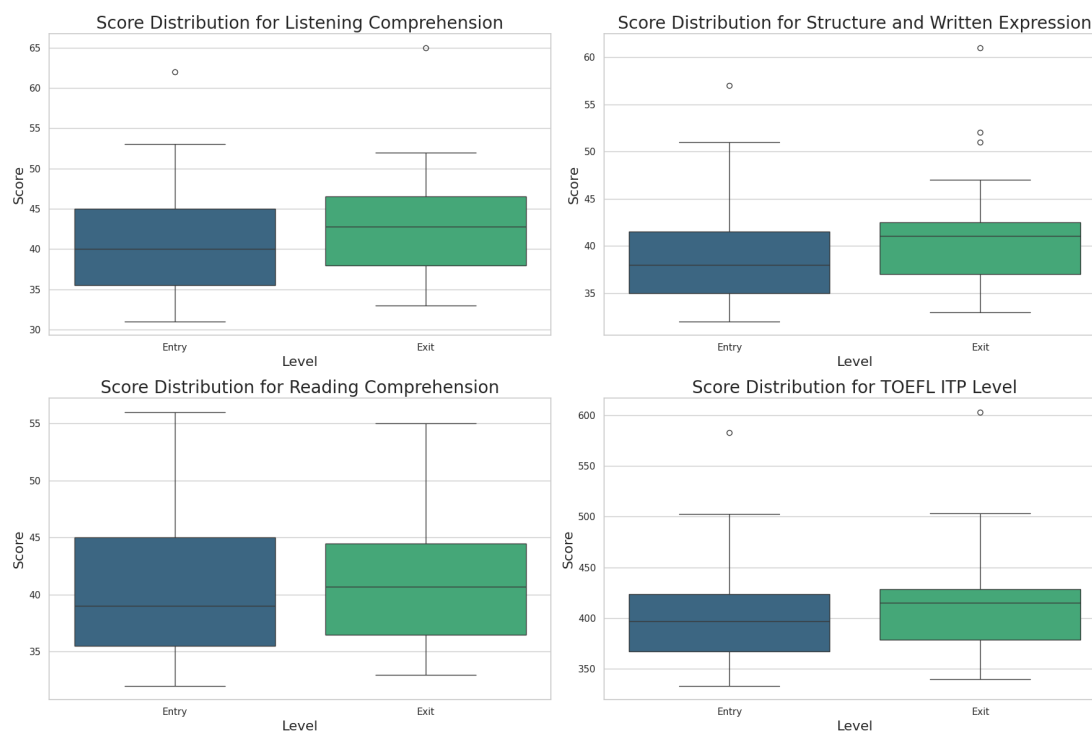


Figure 1: Box plot comparison of entry vs. exit TOEFL ITP scores

Note: For several components, the boxplots suggest that while mean scores increased, the spread (interquartile range) of the Exit scores might have slightly shifted or changed compared to the Entry scores. This means that not all students improved at the same rate, and the distribution of skills at the exit might be different than at entry.

In addition, the correlational analysis depicted in Figure 3 reveals a strong positive relationship between entry and exit scores ($r = 0.7635$), indicating that students' initial proficiency levels are closely associated with their final performance. This suggests that, although improvement is observed, student progress remains consistent with their initial level.

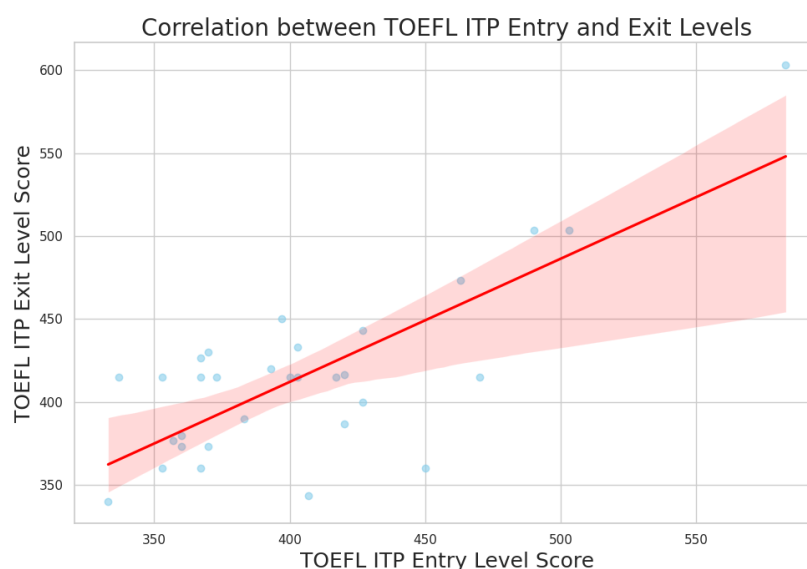


Figure 2: Scatter plot and linear correlation of TOEFL ITP entry vs. exit results

Note: The scatter plot with the trend line shows a positive trend. The points are clustered relatively close to the regression line, which exhibits an upward slope. This visually demonstrates that as the entry score (X-axis) increases, the exit score (Y-axis) also tends to increase.

To evaluate whether these observed differences were statistically significant, a one-tailed t-test was conducted, given that the study focuses on improvement. The null hypothesis (H_0) established that $\mu_{\text{exit}} \leq \mu_{\text{entry}}$, while the alternative hypothesis (H_1) proposed that $\mu_{\text{exit}} > \mu_{\text{entry}}$. A significance level of $\alpha = 0.05$ was applied. The results of the hypothesis test are presented in Table 2.

Table 2: Statistical significance results of the intervention

Parameter	Value	Observation
T-statistic	1.6652	This value represents the magnitude of the difference between the sample means relative to the variation within the samples. A positive value suggests that the Exit Level mean is higher than the Entry Level mean.
P-value (one-sided)	0.0531	This is the probability of observing a T-statistic as extreme as, or more extreme than, the one calculated, assuming the null hypothesis ($\mu_{\text{Exit}} \leq \mu_{\text{Entry}}$) is true. It indicates the strength of evidence against the null hypothesis.
Significance Level (α)	0.05	This is the threshold for statistical significance. We compare the p-value against this level to make a decision about the null hypothesis. An alpha of 0.05 means we are willing to accept a 5% chance of rejecting a true null hypothesis (Type I error).
Decision	Fail to Reject H_0	Since the one-sided p-value (0.0531) is greater than the significance level (0.05), we do not have enough statistical evidence to reject the null hypothesis. <i>"The intervention with IA had no positive effect on the scores"</i> .

Note: The test results indicate that there is no statistically significant evidence at the 0.05 level to suggest that the mean TOEFL ITP score at the Exit Level is higher than at the Entry Level. The observed difference in means could be due to random chance. This does not necessarily mean there was no improvement, but that any improvement was not large enough to be statistically distinguishable from chance at our chosen alpha level.

Although descriptive and correlational analyses show a general and consistent improvement in student performance, hypothesis testing indicates that this improvement is not statistically significant at the 95% confidence level and could therefore be attributed to chance. Consequently, despite the positive trend, further qualitative analysis is needed to better understand the impact of the intervention.

4.2 Students' Perceptions of ChatGPT Use

RQ2: How do students perceive the usefulness, limitations, and academic ethics of using ChatGPT?

The thematic analysis of the focus group identified four main themes related to students' experiences with the integration of ChatGPT: prior familiarity with AI, perceived support for language learning, learning practices with AI, and learner autonomy, as well as perceived limitations and risks.

4.2.1 Prior Familiarity with AI

Participants reported varying levels of experience with ChatGPT before the course. Some had already been introduced to the tool in previous academic contexts. For instance, one participant noted, *"In high school they taught us how to use it to improve our English"* (Noah). Others indicated that their first exposure occurred through peer interaction at university, as reflected in the statement, *"Amelia showed me how to use it"* (Emma). These comments show that students entered the course with varying degrees of familiarity with AI tools.

4.2.2 Perceived Support for Language Learning

The use of ChatGPT as a support tool for language learning emerged as a central theme. Participants reported using the tool for grammar explanations, vocabulary development, and practice exercises. For example, one student explained that *"If I ask it to explain grammar in detail, it becomes easier to understand"* (Emma). Similarly, another participant described using the tool to analyze vocabulary in context: *"I can give it ten words and ask for the definition, what type of word it is, and how to use it in a sentence"* (Amelia). Additionally, students reported using ChatGPT to generate practice materials, as illustrated by the comment, *"I ask it to create exercises for me to practice"* (Olivia and Noah). The results show that ChatGPT was perceived as an important supplementary resource for language learning.

4.2.3 Learning Practices with AI

Participants described incorporating ChatGPT into their study practices, particularly as a tool for clarification and idea generation. One participant stated, *"It helps me clarify the topic more"* (Olivia), while another emphasized efficiency, noting that *"It saved time from looking things up in the dictionary"* (Amelia). Students also viewed ChatGPT as a starting point for generating ideas, as reflected in the statement, *"It is not bad to look for ideas in AI"* (Olivia). These responses highlight ChatGPT's role as a supplementary reference tool integrated into traditional learning practices.

4.2.4 Perceived Limitations and Risks

Despite its benefits, participants also identified important limitations. Some students noted that the information generated by ChatGPT is not always reliable, as reflected in the statement, *"Sometimes the answer is not correct"* (Noah). This led to an awareness of the need for verification, with one participant emphasizing that *"You have to check if the information is correct"* (Noah). Additionally, concerns about overreliance on AI were expressed. One participant stated, *"If ChatGPT does everything, you do not produce the language"* (Amelia). The students' reflections highlight their critical awareness of both the benefits and the risks associated with the use of AI.

4.3 Evidence of Learner Autonomy

RQ3: What evidence of autonomy appears, and how is it related to the use of ChatGPT?

The analysis also revealed evidence of learner autonomy in students' use of ChatGPT. Participants described using the tool to support independent learning, particularly when they encountered difficulties or needed additional resources.

Some students reported using ChatGPT to clarify concepts they did not understand. For example, one participant explained, *"I use it to understand topics I do not understand"* (Noah). Others described using the tool to identify further learning materials, as reflected in the statement, *"It gives me websites or books to learn"* (Olivia). In addition, students demonstrated selective and strategic use of ChatGPT. One participant noted, *"I only use it when I do not understand something"* (Emma), suggesting a conscious decision to rely on the tool only when necessary. The responses provided suggest that ChatGPT can promote learner autonomy by enabling independent problem-solving, access to resources, and self-directed learning strategies.

5. Discussion

The findings of this study provide a nuanced perspective on the role played by artificial intelligence—specifically ChatGPT in supporting the learning of English as a foreign language (EFL) in higher education.

From a quantitative perspective, the results indicate a consistent upward trend in students' TOEFL ITP scores across all components. Although these improvements did not reach statistical significance, they suggest a positive tendency in language development. This finding can be interpreted in light of previous research highlighting the potential of artificial intelligence tools to enhance language learning through personalized feedback, increased practice opportunities, and adaptive learning environments (AlTwijri and Alghizzi, 2024; Bounif, 2026). However, the lack of statistical significance underscores the complexity of measuring the impact of AI-based interventions, suggesting that factors such as the duration of exposure, intensity of use, and learners' initial proficiency levels may mediate these outcomes.

The qualitative findings provide additional insights that help contextualize these results. Students perceive ChatGPT as a valuable tool for improving grammar, vocabulary, and idea generation, which is consistent with prior studies on AI in education (Yang, 2025; Hacıyeva, 2025). However, these findings are interpreted within the framework of the present study and do not reproduce the authors' original wording. At the same time, participants demonstrated a critical awareness of the tool's limitations, particularly regarding the reliability of generated content. This perspective aligns with concerns raised in the literature about the risks of misinformation and the need for guided use of AI technologies in educational contexts (Vieriu and Petrea, 2025).

Importantly, the findings also suggest that ChatGPT may contribute to the development of learner autonomy. Students reported engaging in self-directed practices such as clarifying doubts, seeking additional resources, and using the tool selectively according to their needs. These behaviors reflect key dimensions of learner autonomy as conceptualized by Holec (1981) and Little (1991), particularly in terms of learners' ability to take responsibility

for their learning and regulate their strategies. In this sense, ChatGPT appears to function as a mediating tool that supports autonomous learning practices rather than replacing instructional processes.

The results of this study indicate that ChatGPT's educational value lies not only in its capacity to support language development but also in its potential to foster reflective and self-regulated learning. However, consistent with previous research (Holmes, Bialik, & Fadel, 2019), these benefits depend on its integration into structured learning environments and the presence of pedagogical guidance to ensure its ethical and effective use.

6. Conclusion and Future Work

This study explored the use of ChatGPT as an AI-assisted tool to support English as a Foreign Language (EFL) learning in higher education. Beyond the specific results obtained, the findings contribute to the ongoing discussion on the pedagogical role of generative AI in language education, particularly in contexts where empirical evidence remains limited. The study highlights the relevance of integrating AI tools within structured instructional frameworks, where their use can be aligned with learning objectives and guided by pedagogical criteria.

In this sense, the contribution of ChatGPT should be understood not merely in terms of performance outcomes, but as part of a broader shift toward more flexible, student-centered learning environments. Its use reflects emerging patterns of interaction between learners and technology, where access to immediate support, personalized input, and self-directed exploration plays an increasingly important role in the learning process.

This study has some limitations, including the short duration of the intervention and the specific context of the sample, which may limit the generalizability of the findings. Future research should focus on examining the long-term integration of AI tools in language learning, as well as the conditions under which they can be most effectively incorporated into formal educational settings. In particular, further studies are needed to better understand how pedagogical design, teacher mediation, and contextual factors shape the impact of AI on both language development and learner autonomy.

Ethics Declaration

This study followed the ethical guidelines of Universidad Gerardo Barrios (UGB) and the principles of the Declaration of Helsinki, with informed consent obtained from all participants.

Data were anonymized, securely stored, and participants were free to withdraw at any time without penalty.

AI Declaration

This paper was developed with the support of generative AI tools for language editing and refinement. All ideas, data analysis, and interpretations are the authors' own. AI was not used to generate data or replace critical academic judgment.

Acknowledgments

The authors gratefully acknowledge Alfredo Chévez (lchevez@ugb.edu.sv) for granting permission to conduct this study with his students and for planning the necessary resources for the experimental group. The authors also extend sincere appreciation to the students who participated in this research for their valuable collaboration.

References

- Al-kassab, M. (2022) 'The use of one sample t-Test in the real data', *Journal of advances in mathematics*, 21, pp. 134–138. doi:10.24297/jam.v21i.9279.
- Allen, G. (2020) Understanding AI Technology. Available at: <https://apps.dtic.mil/sti/tr/pdf/AD1099286.pdf> (Accessed: 1 March 2026).
- AlTwijri, L. and Alghizzi, T.M. (2024) 'Investigating the integration of artificial intelligence in English as foreign language classes for enhancing learners' affective factors: A systematic review', *Heliyon*, 10(10), p. e31053. doi:10.1016/j.heliyon.2024.e31053.
- Arquez Muñoz, D.A., Tobón, M.F. and Cely Betancourt, B.L. (2025) 'The advantages of using English as the medium of instruction for globally oriented professional training. A theoretical review', *Ciencia Latina Revista Científica Multidisciplinar*, 9(4), pp. 8376–8398. doi:10.37811/cl_rcm.v9i4.19422.

- Atrianus Toni (2026) 'Challenges faced by EFL students in learning english as a foreign language', *Pedagogic Research-Applied Literacy Journal*, 3(1), pp. 28–33. doi:10.70574/hshvd769.
- Benson, P. (2011) *Teaching and Researching Autonomy*. 2nd edn. London: Routledge. Available at: <https://www.routledge.com/Teaching-and-Researching-Autonomy-in-Language-Learning/Benson/p/book/9781408205013> (Accessed: 18 March 2026).
- Bounif, H. (2026) 'Beneficios de la inteligencia artificial en el aprendizaje de lenguas extranjeras: revisión sistemática', *Horizontes. Revista de Investigación en Ciencias de la Educación*, 10(41), pp. 235–258. doi:10.33996/revistahorizontes.v10i41.1206.
- Braun, V. and Clarke, V. (2006) 'Using thematic analysis in psychology', *Qualitative Research in Psychology*, 3(2), pp. 77–101. doi:10.1191/1478088706qp063oa.
- Education First (2025) EF EPI EF english proficiency index. Available at: <https://www.ef.com/assetscdn/WIBlwq6RdJvcD9bc8RMd/cefcom-epi-site/reports/2025/ef-epi-2025-english.pdf> (Accessed: 17 March 2026).
- ETS (2026) The TOEFL ITP® Assessment Series. Available at: <https://www.ets.org/toefl/itp.html> (Accessed: 1 March 2026).
- Haciyeva, A. (2025) 'Integration of artificial intelligence into language teaching', *EuroGlobal Journal of Linguistics and Language Education*, 2(5), pp. 99–106. doi:10.69760/egjll.2505008.
- Hernández-Sampieri, R. and Mendoza, C. (2018) 'Metodología de la investigación. Las rutas cuantitativa, cualitativa y mixta', 10(18), p. 95. doi:10.22201/fesc.20072236e.2019.10.18.6.
- Holec, H. (1981) *Autonomy and foreign language learning*. Oxford: Pergamon Press.
- Holmes, W., Bialik, M. and Fadel, C. (2019) *Artificial Intelligence in Education. Promise and Implications for Teaching and Learning*. Available at: https://www.researchgate.net/publication/332180327_Artificial_Intelligence_in_Education_Promise_and_Implications_for_Teaching_and_Learning (Accessed: 2 March 2026).
- Kasneji, E. et al. (2023) 'ChatGPT for good? On opportunities and challenges of large language models for education', *Learning and Individual Differences*, 103, p. 102274. doi:10.1016/j.lindif.2023.102274.
- Khan, A., Imran, M. and Javed, M. (2025) 'Language testing and assessment: Validity and reliability in English proficiency exams', *Contemporary journal of social science review*, 03(01), pp. 1950–1959. Available at: https://www.researchgate.net/publication/397488435_LANGUAGE_TESTING_AND_ASSESSMENT_VALIDITY_AND_RELIABILITY_IN_ENGLISH_PROFICIENCY_EXAMS (Accessed: 1 March 2026).
- Kohnke, L., Moorhouse, B.L. and Zou, D. (2023) 'ChatGPT for Language Teaching and Learning', *RELC Journal*, 54(2), pp. 537–550. doi:10.1177/00336882231162868.
- Little, D. (1991) *Learner autonomy. Authentik Language Learning Resources*. Available at: https://www.researchgate.net/publication/259874253_Learner_Autonomy_1_Definitions_Issues_and_Problems (Accessed: 17 March 2026).
- Long, D. and Magerko, B. (2020) 'What is AI literacy? Competencies and design considerations', in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA: ACM, pp. 1–16. doi:10.1145/3313831.3376727.
- Naka, L. and Spahija, D. (2022) 'Impact of English language as a human capital in the Higher Education Institutions' development strategy', *Corporate and Business Strategy Review*, 3(2 special issue), pp. 262–272. doi:10.22495/cbsrv3i2siart7.
- Nuraida, N. (2025) 'Unpacking listening barriers: an investigation into EFL learners' difficulties in the TOEFL ITP context', *GLOBAL: Education Language and Humanity Journal*, 3(1), pp. 23–38. Available at: <https://journal.pcn.ac.id/index.php/global/article/view/98/67> (Accessed: 1 March 2026).
- Okoye, K. and Hosseini, S. (2024) 'Correlation Tests in R: Pearson Cor, Kendall's Tau, and Spearman's Rho', in *R Programming*. Singapore: Springer Nature Singapore, pp. 247–277. doi:10.1007/978-981-97-3385-9_12.
- Pu, P. and Chang, D.Y.-S. (2025) 'Effects of different input modes on blended EFL speaking instruction: a quasi-experimental study', *Computer Assisted Language Learning*, 38(5–6), pp. 1223–1248. doi:10.1080/09588221.2023.2273853.
- Stryker, C. and Kavlakoglu, E. (2025) What is artificial intelligence (AI)? Available at: https://www.ibm.com/think/topics/artificial-intelligence?utm_source=chatgpt.com (Accessed: 1 March 2026).
- Vieriu, A.M. and Petrea, G. (2025) 'The impact of artificial intelligence (AI) on students' academic development', *Education Sciences*, 15(3), p. 343. doi:10.3390/educsci15030343.
- Wijaya, K.F. (2026) 'Challenges in teaching academic writing to EFL learners in higher education contexts', *ELTR Journal*, 10(1), pp. 43–57. doi:10.37147/eltr.v10i1.276.
- Yang, L. (2025) 'The Potential of AI in Education: Personalizing Learning', *International Journal of Artificial Intelligence for Science (IJAI4S)*, 1(2). doi:10.63619/ijai4s.v1i2.003.