

**EVALUATING AI CHATBOTS FOR GRAMMAR AND VOCABULARY ENHANCEMENT
IN EFL ONLINE CLASS: GEMINI AS MINI-TUTOR**

**EVALUACIÓN DE CHATBOTS DE IA PARA LA MEJORA DE GRAMÁTICA Y
VOCABULARIO EN CLASES DE EFL EN LÍNEA: GEMINI COMO MINI-TUTOR**

Autores: ¹Erika Mora Herrera, ²David Gortaire Díaz, ³Gabriela Almache Granda y ⁴Roddy Real Roby.

¹ORCID ID: <https://orcid.org/0000-0002-8156-0557>

²ORCID ID: <https://orcid.org/0000-0001-7364-7305>

²ORCID ID: <https://orcid.org/0000-0003-1858-7121>

⁴ORCID ID: <https://orcid.org/0000-0003-1474-9349>

¹E-mail de contacto: emorah@utb.edu.ec

²E-mail de contacto: dgortaire@utb.edu.ec

³E-mail de contacto: galmache@utb.edu.ec

⁴E-mail de contacto: rreal@utb.edu.ec

Afiliación: ^{1*2*3*4}Universidad Técnica de Babahoyo, (Ecuador).

Artículo recibido: 15 de Febrero del 2026

Artículo revisado: 18 de Febrero del 2026

Artículo aprobado: 23 de Febrero del 2026

¹Licenciada en Lengua y Lingüística Inglesa, titulada en la Universidad de Filosofía Ciencia y Letras de la Educación-Escuela de Lengua, (Ecuador). Maestría en Pedagogía del Idioma Inglés como Lengua Extranjera de la Universidad Bolivariana del Ecuador, (Ecuador).

²Ingeniero en Negocios Internacionales, Escuela Superior Politécnica del Litoral, (Ecuador), 12 años de experiencia laboral. Master en Desarrollo Rural de la Escuela Superior Politécnica del Litoral, (Ecuador), y Master en Pedagogía del Inglés como Lengua Extranjera de la Universidad Bolivariana del Ecuador, (Ecuador).

³Ingeniera en Ciencias Empresariales, mención en Dirección y Planificación Comercial, Universidad de Especialidades Espíritu Santo, (Ecuador). Magíster en Pedagogía de los Idiomas Nacionales y Extranjeros, mención en Enseñanza de Inglés, Universidad Casa Grande, (Ecuador). Magíster en Educación con mención en Innovaciones Pedagógicas, Universidad Casa Grande, (Ecuador). Doctorante en Educación e Innovación, Universidad de Investigación e Innovación de México, (México).

⁴Ingeniero en Negocios Internacionales, titulado por la Escuela Superior Politécnica del Litoral (ESPOL), (Ecuador). Máster Universitario en Dirección de Empresas con mención en Negocios Internacionales por la Universidad de Palermo, (Argentina). Además, posee una Maestría en Pedagogía del Idioma Inglés como Lengua Extranjera de la Universidad Bolivariana del Ecuador, (Ecuador).

Resumen

Esta investigación examina la efectividad de Google Gemini como tutor virtual impulsado por inteligencia artificial para la mejora de gramática y vocabulario en estudiantes de inglés como lengua extranjera (EFL) en entornos de aprendizaje en línea. El estudio empleó un diseño de estudio de caso con métodos mixtos, recopilando datos de percepción de 300 estudiantes de nivel A2.1 del Centro de Idiomas de la Universidad Técnica de Babahoyo, Ecuador, durante un período de intervención de tres meses (octubre-diciembre 2025). Los participantes utilizaron Gemini tanto dentro como fuera de las clases sincrónicas en línea mediante prompts estructurados diseñados para práctica gramatical, desarrollo de vocabulario y práctica conversacional. Los datos se recopilaron a través de encuestas que incluían escalas Likert y preguntas abiertas, analizadas mediante

estadística descriptiva y análisis temático. Los resultados revelan que los estudiantes percibieron a Gemini como una herramienta útil, accesible y generalmente efectiva, con puntuaciones medias de percepción en el rango de "de acuerdo" para apoyo al aprendizaje gramatical ($M = 4.01$), desarrollo de vocabulario ($M = 4.02$), facilidad de uso ($M = 4.06$) y satisfacción general ($M = 4.00$). Los estudiantes valoraron particularmente la disponibilidad 24/7, retroalimentación inmediata y el entorno de práctica sin ansiedad. Sin embargo, se identificaron desafíos incluyendo respuestas ocasionalmente inexactas (29.7%), dificultades técnicas (25.3%) y niveles modestos de compromiso afectivo ($M = 3.82$). El análisis comparativo reveló que los usuarios frecuentes y estudiantes con menos experiencia previa en inglés reportaron percepciones significativamente más positivas. Los hallazgos sugieren que los chatbots de IA funcionan mejor como

herramientas complementarias que extienden la instrucción en aula cuando se implementan con andamiaje pedagógico apropiado, prompts estructurados y apoyo continuo del instructor.

Palabras clave: Chatbots de inteligencia artificial, Aprendizaje de Inglés como lengua extranjera, Google Gemini.

Abstract

This research examines the effectiveness of Google Gemini as an AI-powered virtual tutor for grammar and vocabulary enhancement among English as a Foreign Language (EFL) learners in online learning environments. The study employed a mixed-methods case study design, collecting perception data from 300 A2.1-level students at the Language Center of Universidad Técnica de Babahoyo, Ecuador, during a three-month intervention period (October-December 2025). Participants used Gemini both within and outside synchronous online classes through structured prompts designed for grammar practice, vocabulary development, and conversational practice. Data were collected through surveys including Likert scales and open-ended questions, analyzed using descriptive statistics and thematic analysis. Results reveal that students perceived Gemini as a useful, accessible, and generally effective tool, with mean perception scores in the "agree" range for grammar learning support ($M = 4.01$), vocabulary development ($M = 4.02$), ease of use ($M = 4.06$), and overall satisfaction ($M = 4.00$). Students particularly valued 24/7 availability, immediate feedback, and the anxiety-free practice environment. However, challenges were identified including occasionally inaccurate responses (29.7%), technical difficulties (25.3%), and modest levels of affective engagement ($M = 3.82$). Comparative analysis revealed that frequent users and students with less prior English learning experience reported significantly more positive perceptions. Findings suggest that AI chatbots function best as complementary tools extending classroom instruction when implemented with appropriate pedagogical scaffolding, structured prompts, and ongoing

instructor support. The study contributes empirical evidence to the emerging field of AI-mediated language learning while highlighting implementation considerations essential for maximizing educational benefits and addressing current technological limitations.

Keywords: Artificial intelligence chatbots, English as a foreign language learning, Google Gemini.

Sumário

Esta pesquisa examina a eficácia do Google Gemini como um tutor virtual com inteligência artificial para aprimorar a gramática e o vocabulário de alunos de inglês como língua estrangeira (EFL) em ambientes de aprendizagem online. O estudo empregou um delineamento de estudo de caso com métodos mistos, coletando dados de percepção de 300 alunos de nível A2.1 do Centro de Línguas da Universidade Técnica de Babahoyo, Equador, durante um período de intervenção de três meses (outubro a dezembro de 2025). Os participantes utilizaram o Gemini tanto durante quanto fora das aulas online síncronas, por meio de instruções estruturadas elaboradas para prática gramatical, desenvolvimento de vocabulário e prática de conversação. Os dados foram coletados por meio de questionários que incluíam escalas Likert e questões abertas, e analisados utilizando estatística descritiva e análise temática. Os resultados revelam que os alunos perceberam o Gemini como uma ferramenta útil, acessível e, em geral, eficaz, com pontuações médias de percepção na faixa de "concordo" para suporte à aprendizagem gramatical ($M = 4,01$), desenvolvimento de vocabulário ($M = 4,02$), facilidade de uso ($M = 4,06$) e satisfação geral ($M = 4,00$). Os alunos valorizaram particularmente a disponibilidade 24 horas por dia, 7 dias por semana, o feedback imediato e o ambiente de prática livre de ansiedade. No entanto, foram identificados desafios, incluindo respostas ocasionalmente imprecisas (29,7%), dificuldades técnicas (25,3%) e níveis modestos de engajamento afetivo ($M = 3,82$). A análise comparativa revelou que usuários frequentes e alunos com menos experiência prévia em inglês relataram

percepções significativamente mais positivas. Os resultados sugerem que os chatbots de IA funcionam melhor como ferramentas complementares que ampliam o ensino em sala de aula quando implementados com suporte pedagógico adequado, instruções estruturadas e acompanhamento contínuo pelo instrutor.

Palavras-chave: Chatbots de inteligência artificial, Aprendizagem de inglês como língua estrangeira, Google Gemini.

Introduction

The integration of artificial intelligence (AI) in language education has emerged as a transformative force in contemporary pedagogical practices, particularly in the context of English as a Foreign Language (EFL) instruction (Qassrawi et al., 2024). As educational institutions increasingly adopt digital and hybrid learning modalities, the demand for innovative, personalized, and accessible learning tools has intensified (Wulandari & Purnamaningwulan, 2024). AI-powered chatbots represent a promising technological advancement that addresses critical challenges in EFL education, including limited opportunities for authentic interaction, insufficient personalized feedback, and the scalability constraints inherent in traditional classroom settings (Lara et al., 2023a). These intelligent conversational agents have demonstrated potential to serve as virtual tutors, providing learners with immediate, individualized support for developing fundamental language competencies such as grammar accuracy and vocabulary acquisition (Flores, 2024).

Grammar and vocabulary constitute foundational pillars of language proficiency, serving as essential building blocks for effective communication in English (Walker et al., 2020a). For EFL learners, mastering these linguistic components presents considerable

challenges, as they must internalize complex grammatical structures and extensive lexical repertoires while navigating limited exposure to authentic language use outside the classroom environment (Walker et al., 2020b). Traditional instructional approaches often struggle to provide the intensive, individualized practice necessary for sustained improvement, particularly in large-enrollment online courses where instructor-student ratios constrain opportunities for personalized intervention (Puspitasari et al., 2023). The asynchronous nature of many online EFL programs further exacerbates these limitations, as learners may experience delayed feedback and reduced opportunities for spontaneous language practice (Sholikhah & Ningsih, 2023). Consequently, educational researchers and practitioners have increasingly explored technological solutions capable of supplementing human instruction with scalable, responsive learning support systems.

Among the proliferating landscape of generative AI technologies, Google's Gemini has emerged as a sophisticated large language model with multimodal capabilities and advanced natural language processing functions (Sun et al., 2021). Unlike earlier generation chatbots that relied primarily on rule-based or retrieval-based architectures, Gemini leverages transformer-based neural networks trained on extensive corpora, enabling more nuanced understanding of linguistic contexts, more accurate error detection, and more pedagogically appropriate feedback generation (Zhai y Wibowo, 2023). The potential application of Gemini as a "mini-tutor" in EFL contexts warrants systematic investigation, as its capacity to engage in extended dialogues, provide explanatory feedback, generate contextualized examples, and adapt to individual learner needs may address critical

gaps in online language instruction (Hwang & Nurtantyana, 2022). However, despite the growing enthusiasm surrounding AI chatbots in educational contexts, rigorous empirical evaluation of their effectiveness specifically for grammar and vocabulary enhancement in EFL online environments remains limited (Pikhart, 2020).

The purpose of this study is to evaluate the effectiveness of Gemini as an AI-powered mini-tutor for enhancing grammar accuracy and vocabulary development among EFL learners in online classroom settings. This investigation addresses a critical gap in the literature by systematically examining both the pedagogical efficacy and the practical implementation considerations of integrating Gemini into structured EFL curricula. Specifically, this research explores how Gemini's conversational capabilities, error correction mechanisms, and adaptive feedback functions contribute to measurable improvements in learners' grammatical competence and lexical knowledge. Additionally, the study investigates learner perceptions, engagement patterns, and potential challenges associated with AI-mediated language learning, providing insights essential for educators contemplating the integration of such technologies into their instructional designs. By examining these dimensions through a rigorous methodological framework, this research contributes to the evolving discourse on AI in education while offering practical guidance for EFL practitioners navigating the digital transformation of language teaching.

Materials and Methods

This study employed a case study design with a mixed-methods approach to evaluate the effectiveness of Google Gemini as an AI-powered mini-tutor for grammar and

vocabulary enhancement among EFL learners in online classes. Case study methodology was selected as the most appropriate research design due to its capacity to provide in-depth, contextualized examination of contemporary phenomena within real-life settings, particularly when boundaries between phenomenon and context are not clearly evident (Dela y Dela, 2026). The study population comprised 300 students enrolled at Level 3 (A2.1 according to the Common European Framework of Reference for Languages) at the Language Center of Universidad Técnica de Babahoyo in Ecuador during the October-December 2025 academic term. Participants were selected through convenience sampling, a non-probabilistic sampling technique wherein participants are chosen based on their accessibility and availability to the researcher (Lu, 2019). In the context of this study, the selection of Level 3 students was purposeful and theoretically motivated, as A2.1 learners represent an intermediate-low proficiency level where students possess foundational grammatical knowledge and vocabulary but require extensive practice and reinforcement to consolidate emerging competencies (Conde et al., 2022). The demographic profile of participants reflected the typical composition of university-level EFL programs in Ecuador, with students primarily between 18 and 25 years of age enrolled in various undergraduate programs who were fulfilling institutional English language requirements. The Language Center at Universidad Técnica de Babahoyo serves students across multiple academic faculties, creating a heterogeneous participant pool with diverse disciplinary backgrounds.

The intervention period extended across three months (October through December 2025), encompassing a full academic semester during which participants engaged with Google

Gemini both within synchronous online class sessions and independently for asynchronous practice outside scheduled class time. The integration of Gemini into the curriculum followed principles of blended learning design, wherein technology-mediated activities complemented rather than replaced instructor-led instruction, creating a pedagogical ecology that leveraged affordances of both human teaching and AI-powered practice opportunities (Lara et al., 2023b). During synchronous online class sessions conducted via video conferencing platforms, the instructor introduced specific grammatical structures or vocabulary sets through explicit instruction, modeling, and guided practice, then assigned targeted chatbot interactions as follow-up activities for independent practice and consolidation (Celik, 2023).

To facilitate productive interactions between students and Gemini, researchers developed a structured collection of prompts specifically designed to elicit grammar and vocabulary practice aligned with the Level 3 curriculum objectives. These prompts were carefully crafted following best practices in AI prompt engineering for educational purposes, incorporating clear task specifications, appropriate scaffolding, and explicit guidance for the types of linguistic output expected from students. Students were provided with comprehensive orientation sessions at the beginning of the intervention period that explained the educational rationale for using AI chatbots, demonstrated effective interaction strategies, addressed common concerns about data privacy and appropriate use, and established clear expectations for required chatbot engagement both during and outside class time. Data collection centered primarily on student perceptions of the chatbot activities and their engagement with Gemini, assessed

through a comprehensive survey instrument administered at the conclusion of the intervention period in December 2025. The survey was designed following established principles of questionnaire construction in applied linguistics research, incorporating both closed-ended Likert-scale items for quantitative analysis and open-ended questions for qualitative insights.

The survey instrument was developed through an iterative process involving initial item generation based on relevant literature on technology acceptance and language learning attitudes, expert review by experienced EFL instructors and educational technology specialists, and pilot testing with a small group of students not included in the main study to identify problematic items or unclear wording (Sun et al., 2021). The survey was administered electronically through a secure online platform during the final two weeks of the semester, allowing students to complete it at their convenience while ensuring all responses were collected before course completion. Data analysis proceeded through multiple phases employing descriptive statistical techniques appropriate for survey research and exploratory case study investigation. Quantitative data from Likert-scale survey items were analyzed using SPSS (Statistical Package for the Social Sciences) or equivalent statistical software to calculate measures of central tendency (means, medians) and dispersion (standard deviations, ranges) for each survey item and constructed scale. Qualitative data from open-ended survey questions were analyzed using thematic analysis procedures, a flexible method for identifying, analyzing, and reporting patterns within qualitative data.

Results y Discussion

The analysis of survey data from 300 Level 3 EFL students at Universidad Técnica de Babahoyo revealed generally positive perceptions of Google Gemini as an AI-powered mini-tutor for grammar and vocabulary enhancement during the October-December 2025 intervention period.

Table 1. Demographic Characteristics of Participants

Characteristic	Category	n	%
Gender	Female	178	59.3
	Male	118	39.3
	Non-binary/Prefer not to say	4	1.3
Age	18-20 years	162	54.0
	21-23 years	98	32.7
	24-26 years	32	10.7
	27+ years	8	2.7
Academic Faculty	Engineering	82	27.3
	Business Administration	68	22.7
	Health Sciences	54	18.0
	Education	47	15.7
	Agricultural Sciences	35	11.7
Prior English Learning Experience	Less than 2 years	45	15.0
	2-4 years	128	42.7
	5-7 years	97	32.3
	More than 7 years	30	10.0
Previous AI Chatbot Experience	None	187	62.3
	Limited (1-3 times)	76	25.3
	Moderate (4-10 times)	28	9.3
	Extensive (10+ times)	9	3.0
Daily Internet Access	Limited (less than 2 hours)	23	7.7
	Moderate (2-4 hours)	89	29.7
	Substantial (5-8 hours)	142	47.3
	Extensive (8+ hours)	46	15.3

Source: Own elaboration

Table 1 presents the demographic composition of the 300 participants who completed the study. The sample demonstrated considerable diversity across multiple characteristics relevant to interpretation of findings. Gender distribution showed a moderate female majority (59.3%), which is consistent with enrollment patterns in language programs at Ecuadorian universities. The age distribution revealed that the majority of participants were traditional university-age students, with 86.7% falling between 18 and 23 years old, though a smaller proportion of mature students (13.3% aged 24 or older) provided perspectives from learners balancing academic study with work or family responsibilities. Academic faculty representation spanned six major disciplinary areas, with Engineering (27.3%) and Business

Administration (22.7%) constituting the largest groups, reflecting the institutional profile of Universidad Técnica de Babahoyo as a comprehensive university with strengths in applied and professional programs.

Table 2. Descriptive Statistics for Student Perceptions of Gemini for Grammar Learning

Item	M	SD	Median	Mode	Min	Max
Gemini helped me understand English grammar rules better	4.12	0.78	4.00	4	2	5
The grammar explanations provided by Gemini were clear and easy to understand	3.98	0.83	4.00	4	1	5
Gemini accurately identified grammar errors in my writing	3.85	0.91	4.00	4	1	5
The feedback on grammar from Gemini was helpful for improving my English	4.05	0.81	4.00	4	2	5
I could practice grammar structures more frequently using Gemini than in regular classes	4.23	0.76	4.00	5	2	5
Gemini provided examples that helped me understand how to use grammar correctly	4.08	0.79	4.00	4	2	5
Using Gemini increased my confidence in using English grammar	3.92	0.88	4.00	4	1	5
The grammar corrections from Gemini were appropriate to my level (A2.1)	3.87	0.86	4.00	4	1	5
Overall Grammar Learning Perception Scale	4.01	0.68	4.06	—	2.13	5.00

Source: Own elaboration

Table 2 displays descriptive statistics for eight items assessing student perceptions of Gemini's effectiveness for grammar learning, along with the composite Grammar Learning Perception Scale. The overall scale mean of 4.01 (SD = 0.68) indicates that participants generally agreed that Gemini provided valuable support for grammar development, with the composite score falling solidly in the "agree" range and demonstrating acceptable internal consistency ($\alpha = .89$). Among individual items, participants expressed strongest agreement with the statement "I could practice grammar structures more frequently using Gemini than in regular classes" (M = 4.23, SD = 0.76), highlighting a key affordance of AI chatbots identified in previous research: their capacity to provide unlimited practice opportunities beyond the temporal and logistical constraints of traditional instruction (Pikhart, 2020). Students also positively evaluated Gemini's explanatory capabilities, with mean scores above 4.0 for items related to understanding grammar rules

better, receiving helpful feedback, and obtaining useful examples of correct usage.

Table 3. Descriptive Statistics for Student Perceptions of Gemini for Vocabulary Development

Item	M	SD	Median	Mode	Min	Max
Gemini helped me learn new English vocabulary words	4.18	0.74	4.00	4	2	5
The vocabulary explanations from Gemini were clear and included good examples	4.06	0.79	4.00	4	2	5
Gemini helped me understand how to use vocabulary words in different contexts	3.95	0.84	4.00	4	1	5
I learned more vocabulary using Gemini than I would have learned without it	4.02	0.82	4.00	4	1	5
Gemini provided useful synonyms and related words that expanded my vocabulary	4.15	0.77	4.00	4	2	5
The chatbot helped me remember vocabulary words better through conversation practice	3.88	0.89	4.00	4	1	5
Using Gemini increased my confidence in using new vocabulary words	3.91	0.86	4.00	4	1	5
Gemini helped me understand the differences between similar vocabulary words	3.97	0.83	4.00	4	1	5
Overall Vocabulary Development Perception Scale	4.02	0.66	4.00	—	2.25	5.00

Source: Own elaboration

Table 3 presents descriptive statistics for items measuring student perceptions of Gemini's contribution to vocabulary development. The overall Vocabulary Development Perception Scale achieved a mean of 4.02 (SD = 0.66) with excellent internal consistency ($\alpha = .91$), indicating strong agreement that the chatbot effectively supported lexical learning. Students rated highest the item "Gemini helped me learn new English vocabulary words" (M = 4.18, SD = 0.74), suggesting that participants perceived genuine vocabulary gains from their chatbot interactions. Additional highly rated aspects included Gemini's provision of useful synonyms and related words (M = 4.15, SD = 0.77) and clear explanations with good examples (M = 4.06, SD = 0.79). Overall, these results indicate that Gemini was perceived as an effective tool for vocabulary expansion and initial learning, though its contribution to deeper lexical processing and long-term retention may have been more variable across students.

Table 4. Descriptive Statistics for Ease of Use, Engagement, and Satisfaction with Gemini (N = 300)

Dimension/Item	M	SD	Median	Mode
Ease of Use				
Gemini was easy to access and use	4.32	0.72	4.00	5
I did not experience technical problems when using Gemini	3.67	1.04	4.00	4
The chatbot interface was user-friendly and intuitive	4.08	0.81	4.00	4
I could easily understand how to interact with Gemini	4.15	0.76	4.00	4
Ease of Use Scale	4.06	0.69	4.00	—
Engagement				
Interacting with Gemini was interesting and engaging	3.89	0.87	4.00	4
I enjoyed practicing English with the chatbot	3.76	0.94	4.00	4
Using Gemini motivated me to practice English more frequently	3.82	0.91	4.00	4
I felt more comfortable making mistakes with Gemini than with my teacher or classmates	4.21	0.79	4.00	5
I looked forward to chatbot practice activities	3.54	0.98	4.00	4
The chatbot kept my attention during practice sessions	3.68	0.93	4.00	4
Engagement Scale	3.82	0.74	3.83	—
Overall Satisfaction				
Overall, I am satisfied with using Gemini for learning English	3.94	0.84	4.00	4
I would recommend using Gemini to other English students	3.98	0.82	4.00	4
I plan to continue using AI chatbots for language learning in the future	4.05	0.80	4.00	4
Using Gemini was a valuable addition to my English course	4.02	0.79	4.00	4
Overall Satisfaction Scale	4.00	0.72	4.00	—

Source: Own elaboration

Table 4 presents three conceptually distinct but related dimensions of student perceptions: ease of use, engagement, and overall satisfaction with Gemini. The Ease of Use scale achieved a mean of 4.06 (SD = 0.69, $\alpha = .84$), indicating that students generally found Gemini accessible and user-friendly. The highest-rated ease of use item was "Gemini was easy to access and use" (M = 4.32, SD = 0.72), suggesting that the technical implementation was successful in minimizing barriers to student participation. However, the item "I did not experience technical problems when using Gemini" received notably lower ratings (M = 3.67, SD = 1.04) and exhibited the highest standard deviation in the entire survey, indicating substantial variability in technical experiences, suggesting that despite modest engagement levels, the perceived benefits were sufficient to motivate continued adoption beyond the required intervention period.

Table 5. Frequency of Gemini Usage Patterns During the Intervention Period

Usage Pattern	n	%
Frequency of Use (per week)		
Rarely (1-2 times)	48	16.0

Occasionally (3-4 times)	127	42.3
Frequently (5-6 times)	89	29.7
Very frequently (7+ times)	36	12.0
Average Duration per Session		
Short (less than 10 minutes)	62	20.7
Moderate (10-20 minutes)	152	50.7
Extended (21-30 minutes)	64	21.3
Very extended (more than 30 minutes)	22	7.3
Primary Context of Use		
Only during class activities	56	18.7
Mostly during class, occasionally outside	98	32.7
Equally during and outside class	87	29.0
Mostly outside class, occasionally during	47	15.7
Only outside class for independent practice	12	4.0
Most Common Activity Types (multiple selections allowed)		
Grammar error correction practice	234	78.0
Vocabulary expansion and definitions	267	89.0
Conversational practice on assigned topics	189	63.0
Requesting explanations of grammar rules	212	70.7
Getting examples of vocabulary in context	245	81.7
Practicing writing sentences/paragraphs	156	52.0
Asking questions about English language	198	66.0
Device Used for Access		
Smartphone	203	67.7
Laptop computer	71	23.7
Desktop computer	18	6.0
Tablet	8	2.7

Source: Own elaboration

Table 5 provides detailed information about how students actually engaged with Gemini during the three-month intervention period, revealing considerable variability in usage patterns across the participant sample. Regarding frequency of use, the modal pattern was occasional use (3-4 times per week, 42.3%), with substantial minorities engaging frequently (5-6 times weekly, 29.7%) or very frequently (7+ times weekly, 12.0%), while 16.0% used the chatbot rarely (1-2 times weekly). Regarding context of use, the data revealed relatively balanced distribution between primarily in-class use (51.4% when combining the first two categories) and substantial out-of-class independent practice (48.7% when combining the last three categories), suggesting successful implementation of the blended learning design where chatbot activities bridged formal instruction and autonomous learning. The activity type data provides valuable insights into how students actually utilized Gemini, with

vocabulary expansion and definitions being the most common activity (89.0%), followed closely by getting examples of vocabulary in context (81.7%) and grammar error correction practice (78.0%).

Table 6. Comparison of Perception Scores Across Student Characteristics

Characteristic	Grammar Learning M (SD)	Vocabulary Development M (SD)	Engagement M (SD)	Overall Satisfaction M (SD)
Gender				
Female (n = 178)	4.08 (0.65)	4.09 (0.62)	3.87 (0.71)	4.05 (0.69)
Male (n = 118)	3.91 (0.72)	3.92 (0.71)	3.74 (0.78)	3.92 (0.76)
Age Group				
18-20 years (n = 162)	4.05 (0.67)	4.06 (0.65)	3.91 (0.72)	4.08 (0.70)
21-23 years (n = 98)	3.98 (0.69)	3.99 (0.68)	3.76 (0.75)	3.94 (0.73)
24+ years (n = 40)	3.93 (0.71)	3.94 (0.67)	3.65 (0.79)	3.82 (0.78)
Prior English Learning Experience				
Less than 4 years (n = 173)	4.09 (0.66)	4.11 (0.63)	3.89 (0.71)	4.07 (0.69)
5+ years (n = 127)	3.90 (0.70)	3.89 (0.69)	3.71 (0.77)	3.90 (0.75)
Previous AI Chatbot Experience				
None (n = 187)	3.96 (0.70)	3.98 (0.68)	3.76 (0.76)	3.94 (0.74)
Some experience (n = 113)	4.11 (0.63)	4.09 (0.62)	3.93 (0.69)	4.10 (0.67)
Frequency of Gemini Use				
Rarely/Occasionally	3.87 (0.72)	3.89 (0.71)	3.65 (0.79)	3.82 (0.77)
Frequently/Very frequently	4.21 (0.58)	4.21 (0.56)	4.06 (0.62)	4.24 (0.61)
Academic Faculty				
STEM (Engineering, Health, Agriculture) (n = 171)	3.97 (0.69)	3.99 (0.67)	3.79 (0.75)	3.96 (0.73)
Non-STEM (Business, Education, Social Sciences) (n = 129)	4.07 (0.66)	4.06 (0.65)	3.85 (0.72)	4.05 (0.70)
Overall Sample (N = 300)	4.01 (0.68)	4.02 (0.66)	3.82 (0.74)	4.00 (0.72)

Source: Own elaboration

Table 7 presents comparative descriptive statistics examining whether perception scores varied systematically across different student characteristics, providing insights into which learner populations may benefit most from AI chatbot integration and which may require additional support or modified approaches. Gender comparisons revealed that female students reported slightly higher mean scores across all perception dimensions, with the largest difference observed for grammar learning perceptions (Female M = 4.08 vs. Male M = 3.91). While these differences appear modest in absolute terms, they may reflect documented gender differences in technology acceptance, learning strategy preferences, or willingness to engage with novel educational technologies, though interpretations must be

cautious given the descriptive nature of these comparisons (Tarhini et al., 2017).

Age group comparisons showed a pattern wherein younger students (18-20 years) consistently reported more positive perceptions than older students (24+ years), with the largest difference again appearing in engagement scores (18-20 years $M = 3.91$ vs. 24+ years $M = 3.65$). This finding might reflect greater digital nativity and comfort with AI technologies among younger learners, or alternatively, could indicate that mature students with established learning preferences found the chatbot less compatible with their preferred approaches to language study. The comparison by prior English learning experience revealed an unexpected pattern: students with less experience (under 4 years) reported more positive perceptions across all dimensions than those with more extensive backgrounds (5+ years).

The findings from this case study provide empirical evidence that Google Gemini can serve as an effective AI-powered mini-tutor for grammar and vocabulary enhancement in EFL online classes, while also revealing important nuances, limitations, and implementation considerations that warrant careful attention from educators and researchers. The overall pattern of results indicates that students at the A2.1 proficiency level perceived Gemini as a useful, accessible, and generally effective tool for language learning, with mean perception scores consistently falling in the "agree" range across dimensions of grammar learning support, vocabulary development, ease of use, and overall satisfaction.

These quantitative findings align with emerging research suggesting that contemporary large language models possess sufficient linguistic

sophistication and pedagogical capabilities to provide meaningful support for language learners, particularly in domains requiring extensive practice, immediate feedback, and personalized explanations (Wang, 2019; Wang, 2022). However, the qualitative data and variability observed across perception items reveal a more complex picture wherein the effectiveness of chatbot-mediated learning depended critically on factors such as implementation design, student engagement patterns, individual learning preferences, and the specific linguistic features being practiced.

Students' positive perceptions of Gemini's contributions to grammar learning ($M = 4.01$, $SD = 0.68$) and vocabulary development ($M = 4.02$, $SD = 0.66$) suggest that the chatbot successfully addressed key challenges inherent in EFL instruction, particularly the provision of extensive practice opportunities and immediate, personalized feedback that are difficult to deliver consistently in traditional classroom settings (Astrini et al., 2024). The finding that students most strongly endorsed the statement about practicing grammar structures more frequently with Gemini than in regular classes ($M = 4.23$) empirically validates theoretical claims that AI chatbots' primary pedagogical value lies in their capacity to supplement rather than replace human instruction by providing unlimited, on-demand practice that extends learning beyond temporal and spatial constraints of formal lessons (Bibauw et al., 2019; Huang et al., 2022).

For vocabulary development, the strong ratings for Gemini's provision of synonyms, related words, and contextual examples ($M = 4.06$ - 4.18 across relevant items) suggest alignment with research-supported vocabulary teaching principles emphasizing semantic elaboration, network building, and encounter with words in

varied authentic contexts (Boers, 2021; Lindner et al., 2019). However, the relatively lower rating for the item assessing retention through conversation practice ($M = 3.88$, $SD = 0.89$) and the qualitative observation that vocabulary gains may have been concentrated in receptive knowledge raises questions about the depth of processing actually occurring during chatbot interactions. This interpretation suggests that effectiveness of chatbots for vocabulary learning may depend critically on task design, with more open-ended communicative activities potentially yielding deeper processing than definitional queries or structured exercises, though this hypothesis requires empirical investigation through comparative studies examining different interaction types.

The variability observed across student experiences and the identification of both significant advantages and challenges point to the critical importance of thoughtful pedagogical design and implementation support rather than simply deploying technology and expecting positive outcomes (Chen et al., 2020; Wang, 2022). The TPACK framework emphasizes that effective technology integration requires intersection of technological knowledge, pedagogical knowledge, and content knowledge, and findings from this study reinforce that providing students with chatbot access alone is insufficient—success requires structured guidance on how to interact productively, explicit training in prompt formulation, integration with curricular objectives and classroom instruction, and ongoing instructor support to troubleshoot problems and mediate learning (Celik, 2023; Reyes et al., 2017).

The instructor's role in chatbot-enhanced courses emerges as fundamentally important but qualitatively different from traditional

teaching roles. Rather than serving as the primary source of input, explanation, and feedback, instructors in chatbot-integrated contexts function more as designers of learning sequences, curators of productive prompts and activities, monitors of student engagement patterns, interveners when students encounter difficulties or confusion, and facilitators of metalinguistic reflection that helps students extract generalizable knowledge from their chatbot interactions (Lindner et al., 2019; Wulandari y Purnamaningwulan, 2024). This study's findings contribute to evolving theoretical understanding of how AI chatbots function within language learning ecologies and what mechanisms account for their pedagogical effects (Hernández & Rodríguez, 2024; Lu, 2019). The findings also have implications for understanding the role of feedback in language learning. The immediate, individualized feedback provided by Gemini represents a form of automated written corrective feedback that has been extensively studied in CALL research, but with important distinctions from traditional automated feedback systems (Divekar* et al., 2022). The chatbot context may influence uptake in complex ways: reduced anxiety might increase receptiveness to feedback, but the non-human source might reduce credibility or attention compared to feedback from respected human instructors (Lameras & Arnab, 2021).

Conclusions

This case study investigated the effectiveness of Google Gemini as an AI-powered mini-tutor for grammar and vocabulary enhancement among 300 A2.1-level EFL students at Universidad Técnica de Babahoyo during a three-month online course intervention. The findings reveal that students generally perceived Gemini as a valuable, accessible, and effective learning tool, with mean perception scores consistently in the "agree" range across dimensions of grammar

learning support ($M = 4.01$), vocabulary development ($M = 4.02$), ease of use ($M = 4.06$), and overall satisfaction ($M = 4.00$). Qualitative analysis identified key affordances that students valued most highly: 24/7 availability enabling flexible practice schedules, immediate personalized feedback addressing individual learning needs, and a low-anxiety environment facilitating risk-taking and experimentation without fear of social judgment. These advantages address persistent challenges in EFL education, particularly the provision of extensive, individualized practice opportunities that exceed what is feasible in traditional classroom settings constrained by time, resources, and instructor-student ratios (An et al., 2023).

However, the study also revealed important limitations and implementation challenges that temper uncritical enthusiasm about AI chatbots in language education. Students reported occasional inaccurate or confusing responses from Gemini (29.7%), technical connectivity difficulties (25.3%), and concerns about knowing when to trust chatbot feedback (28.0%), highlighting that current LLM technology, while sophisticated, remains imperfect and requires critical user engagement rather than passive acceptance. The modest engagement scores ($M = 3.82$) relative to perceived usefulness suggest that while students recognized the instrumental value of chatbot practice, many did not find interactions intrinsically motivating or enjoyable, with usage patterns showing considerable variability and only 12% of students achieving daily practice. These findings underscore that successful chatbot integration requires thoughtful pedagogical design including structured prompts, explicit training in effective usage, integration with course requirements and assessments, ongoing instructor support, and

deliberate strategies to sustain engagement beyond initial novelty (Godwin, 2022; Pérez, 2021).

The theoretical and practical contributions of this research include empirical documentation of student perceptions across multiple dimensions of chatbot-mediated learning in an authentic educational context, identification of specific affordances and limitations of contemporary LLMs for grammar and vocabulary instruction, development and validation of structured prompt frameworks aligned with curricular objectives, and illumination of implementation factors that influence effectiveness including usage frequency, student characteristics, and pedagogical scaffolding. The finding that high-frequency users reported substantially more positive perceptions than occasional users points to the critical importance of promoting sustained, consistent engagement rather than sporadic use. The comparative analyses suggesting that less experienced learners and students with prior chatbot exposure reported more favorable perceptions indicate that effectiveness may vary across learner populations and that targeted support may be needed for skeptical or technologically inexperienced students.

Important limitations constrain the conclusions that can be drawn from this study, most notably the reliance on self-reported perception data without direct measures of learning outcomes, the case study design precluding causal inference, and the convenience sampling limiting generalizability beyond the specific Ecuadorian university context. Future research should address these limitations through experimental or quasi-experimental designs with comparison groups, incorporation of validated pre-post assessments measuring

actual grammar and vocabulary gains, analysis of chatbot interaction transcripts to understand productive versus unproductive engagement patterns, and longitudinal investigations examining whether benefits accumulate and engagement sustains over extended timeframes (Mackey & Gass, 2016; Plonsky & Oswald, 2014). Additionally, research is needed on optimal integration strategies including ideal balance between chatbot and human instruction, most effective prompt designs for different learning objectives and proficiency levels, and interventions to support sustained motivation and self-regulated practice.

Referencias Bibliográficas

- An, X., Chai, C., Li, Y., Zhou, Y., Shen, X., Zheng, C., & Chen, M. (2023). Modeling English teachers' behavioral intention to use artificial intelligence in middle schools. *Education and Information Technologies*, 28(5), 5187–5208. <https://doi.org/10.1007/s10639-022-11286-z>
- Astrini, A., Prastiwi, Y., & Sutopo, A. (2024). Exploring students' experience of using Nearpod in grammar lesson in relation with students' increased engagement: A descriptive qualitative approach. *Journal of International Multidisciplinary Research*, 2(5), 205–214.
- Boers, F. (2021). *Evaluating second language vocabulary and grammar instruction: A synthesis of the research on teaching words, phrases, and patterns*. Routledge.
- Celik, I. (2023). Towards Intelligent-TPACK: An empirical study on teachers' professional knowledge to ethically integrate artificial intelligence-based tools into education. *Computers in Human Behavior*, 138, 107468. <https://doi.org/10.1016/j.chb.2022.107468>
- Chen, L., Chen, P., & Lin, Z. (2020). Artificial intelligence in education: A review. *IEEE Access*, 8, 75264–75278. <https://doi.org/10.1109/ACCESS.2020.2988510>
- Conde, L., Cueva, G., Chamba, L., & Ureña, M. (2022). Impact of artificial intelligence in basic general education in Ecuador. *2022 17th Iberian Conference on Information Systems and Technologies (CISTI)*, 1–7.
- Dela, C., & Dela, L. (2026). Leveraging artificial intelligence for intelligent student support: An AI-enabled SRM framework for higher education. *Multidisciplinary Science Journal*, 8(3). <https://doi.org/10.31893/multiscience.2026160>
- Divekar, R., Drozdal, J., Chabot, S., Zhou, Y., Su, H., Chen, Y., Zhu, H., Hendler, J., & Braasch, J. (2022). Foreign language acquisition via artificial intelligence and extended reality: Design and evaluation. *Computer Assisted Language Learning*, 35(9), 2332–2360. <https://doi.org/10.1080/09588221.2021.1879162>
- Flores, C. (2024). La evaluación educativa en la era de la inteligencia artificial: Cambios de paradigmas. *LATAM Revista Latinoamericana de Ciencias Sociales y Humanidades*, 5(1). <https://doi.org/10.56712/latam.v5i1.1694>
- Hernández, N., & Rodríguez, M. (2024). Inteligencia artificial aplicada a la educación y la evaluación educativa en la universidad: Introducción de sistemas de tutorización inteligentes y otras tendencias futuras. *Revista de Educación a Distancia*, 24(78). <https://doi.org/10.6018/red.594651>
- Hwang, W., & Nurtantyana, R. (2022). The integration of multiple recognition technologies and artificial intelligence to facilitate EFL writing in authentic contexts. *International Conference on Information Technology (InCIT)*, 379–383. <https://doi.org/10.1109/InCIT56086.2022.10067490>
- Lameras, P., & Arnab, S. (2021). Power to the teachers: An exploratory review on artificial intelligence in education. *Information*, 13(1), 14.
- Lara, R., Criollo, L., Calderón, C., & Matamba, B. (2023). La inteligencia artificial: Análisis del presente y futuro en la educación

- superior. *Revista Científica Multidisciplinar G-Nerando*, 4(1).
- Pikhart, M. (2020). Intelligent information processing for language education: The use of artificial intelligence in language learning apps. *Procedia Computer Science*, 176, 1412–1419.
<https://doi.org/10.1016/j.procs.2020.09.151>
- Puspitasari, I., Pureka, M., & Azizah, N. (2023). Integrating ICT: EFL students challenge in learning grammar. *Proceeding of Conference on English Language Teaching (CELT I 2023)*, 3, 433–441.
- Qassrawi, R., ElMashharawi, A., Itmeizeh, M., & Tamimi, M. (2024). AI-powered applications for improving EFL students' speaking proficiency in higher education. *Forum for Linguistic Studies*, 6(5), 535–549.
<https://doi.org/10.30564/fls.v6i5.6966>
- Sholikhah, N., & Ningsih, F. (2023). Interactive pathways: Exploring students' acceptance of using Nearpod for English grammar proficiency. *Jurnal Bahasa Lingua Scientia*, 15(2), 415–439.
- Sun, Z., Anbarasan, M., & Kumar, D. (2021). Design of online intelligent English teaching platform based on artificial intelligence techniques. *Computational Intelligence*, 37(3), 1166–1180.
- Walker, N., Monaghan, P., Schoetensack, C., & Rebuschat, P. (2020). Distinctions in the acquisition of vocabulary and grammar: An individual differences approach. *Language Learning*, 70(S2), 221–254.
- Wang, Z. (2022). Computer-assisted EFL writing and evaluations based on artificial intelligence: A case from a college reading and writing course. *Library Hi Tech*, 40(1), 80–97.
- Wulandari, M., & Purnamaningwulan, R. (2024). Exploring Indonesian EFL pre-service teachers' experiences in AI-assisted teaching practicum: Benefits and drawbacks. *LLT Journal: A Journal on Language and Language Teaching*, 27(2), 878–894.
- Zhai, C., & Wibowo, S. (2023). A systematic review on artificial intelligence dialogue systems for enhancing English as foreign language students' interactional competence in the university. *Computers and Education: Artificial Intelligence*, 4, 100134.
<https://doi.org/10.1016/j.caeai.2023.100134>



Esta obra está bajo una licencia de **Creative Commons Reconocimiento-No Comercial 4.0 Internacional**. Copyright © Erika Mora Herrera, David Gortaire Díaz, Gabriela Almache Granda y Roddy Real Roby.

