UNEMI
UNIVERSIDAD ESTATAL DE MILAGRO

REPÚBLICA DEL ECUADOR

UNIVERSIDAD ESTATAL DE MILAGRO

VICERRECTORADO DE INVESTIGACIÓN Y POSGRADO

FACULTAD DE POSGRADOS

ARTÍCULOS PROFESIONALES DE ALTO NIVEL
PREVIO A LA OBTENCIÓN DEL TÍTULO DE:

MAGÍSTER EN MAGÍSTER EN ENSEÑANZA DE INGLÉS COMO LENGUA EXTRANJERA

TEMA:

EFFECTIVENESS OF ARTIFICIAL INTELLIGENCE TOOLS IN THE DEVELOPMENT OF ENGLISH LANGUAGE VOCABULARY: A STUDY WITH SEVENTH GRADE STUDENTS OF BASIC GENERAL EDUCATION

Autor:

DORIS FABIOLA FUENTES RENDÓN

Tutor:

JORGE FRANCISCO ZAMBRANO PACHAY

Milagro, 2026

RESUMEN

El vocabulario es uno de los pilares del aprendizaje del inglés como lengua extranjera, y hoy las herramientas de inteligencia artificial (IA) abren nuevas posibilidades para fortalecerlo en el aula. Este estudio buscó determinar en qué medida el uso de herramientas de IA contribuye al desarrollo del vocabulario del idioma inglés en estudiantes de séptimo año de Educación General Básica. Para ello, se trabajó con un enfoque mixto de diseño explicativo secuencial: la fase cuantitativa siguió un diseño preexperimental de un solo grupo, con mediciones antes y después de la intervención, y participaron 27 estudiantes de entre 11 y 12 años. La intervención duró ocho semanas, con dos sesiones semanales de 40 minutos apoyadas en Duolingo. El vocabulario se midió con una prueba de 20 ítems, analizando los datos mediante la prueba t de Student para muestras relacionadas y el tamaño del efecto d de Cohen. Los resultados mostraron un incremento claro: la media pasó de 10,30 puntos en el pretest a 13,22 en el postest, con una diferencia estadísticamente significativa, $t\,(26) = 7{,}68$; $p < 0{,}001$, y un tamaño del efecto grande (d = 1,48). Desde lo cualitativo, se entendió que la motivación, la práctica constante, la retroalimentación inmediata y el acompañamiento del docente aparecieron como piezas clave para explicar el progreso alcanzado. Por ello, integrar herramientas de IA de forma pedagógicamente planificada tiene potencial real para fortalecer el vocabulario en inglés, sobre todo cuando se emplean como apoyo complementario dentro de un proceso de enseñanza bien estructurado.

PALABRAS CLAVES

Inteligencia artificial; vocabulario en inglés; aprendizaje del inglés; educación general básica; tecnología educativa.

ABSTRACT

Vocabulary is one of the cornerstones of learning English as a foreign language, and today artificial intelligence (AI) tools are opening new possibilities for strengthening it in the classroom. This study sought to determine the extent to which the use of AI tools contributes to the development of English vocabulary in seventh-grade students. A mixed-methods approach with a sequential explanatory design was used: the quantitative phase followed a single-group pre-experimental design, with measurements taken before and after the intervention, and involved 27 students between 11 and 12 years old. The intervention lasted eight weeks, with two 40-minute sessions per week using Duolingo. Vocabulary was measured with a 20-item test, and the data were analyzed using the paired-samples t-test and Cohen's d effect size. The results showed a clear increase: the mean score rose from 10.30 points in the pretest to 13.22 in the posttest, with a statistically significant difference, $t\,(26) = 7.68$. $p < 0.001$, and a large effect size (d = 1.48). From a qualitative perspective, motivation, consistent practice, immediate feedback, and teacher support emerged as key factors in explaining the progress achieved. Therefore, integrating AI tools in a pedagogically planned manner has real potential to strengthen English vocabulary, especially when used as supplementary support within a well-structured teaching process.

KEYWORDS

Artificial intelligence; English vocabulary; English language learning; basic general education; educational technology.

1. INTRODUCCIÓN (OBJETIVO DEL ARTÍCULO)

Just a few years ago, talking about artificial intelligence (AI) in the English classroom sounded almost like science fiction. Today, however, conversational chatbots, speech recognition systems, and adaptive learning platforms are part of the everyday repertoire of many foreign language teachers. The review by Abusahyon et al. (2023) confirms this trend: AI-powered chatbots are now one of the most widely used tools for personalizing instruction and offering interactive practice in second language acquisition. Their arrival has raised a question that remains unanswered: do they really help students learn more and better, or are they just another technological fad? Ayala & Alvarado (2023) place the debate somewhere in between: they acknowledge the potential of these tools to personalize instruction and offer immediate feedback, but caution that their real impact depends on the conditions under which they are implemented. This caveat is especially important in Latin America, where learning English as a foreign language still suffers from fundamental problems: inequality in access to technology, little communicative exposure to the language, and classrooms where new vocabulary is usually learned through repetition and memorization, rather than meaningful use.

Recent literature, however, is beginning to nuance this picture. In their study with B1 level students, Guagchinga et al. (2024) found that the use of AI tools increases both motivation and autonomy in learning, although with one significant condition: everything depends on equitable access and on the teacher knowing how to integrate the tool into what they were already doing in class, not simply replacing it. A similar finding is documented by Chicaíza et al. (2023) , whose research using ChatGPT as a virtual tutor in Ecuadorian schools reports tangible improvements in oral and written production, as well as a vocabulary that expands to higher-than-expected proficiency levels. And when Juarez et al. (2024) examine the same phenomenon in university students, they reach a conclusion worth highlighting: the benefit is not only lexical. Students using ChatGPT also report less anxiety when speaking or writing in English, a side effect rarely mentioned in vocabulary studies, but one that likely plays a role.

But how representative is this evidence? Herein lies the first serious problem. Berdella & Albornoz's (2025) systematic review of the literature published between 2019 and 2023 confirms that Latin American research on AI and English language learning is on the rise, but it also reveals a worrying trend: almost all of this research is concentrated in universities and high schools. Basic general education has simply been left off the radar. Torres (2025) reaches a very similar conclusion in his own review: chatbots and conversational assistants are by far the most studied tools, followed distantly by image generators and role-playing environments, but again, basic education is practically absent from the picture.

Taken together, the picture reveals a gap that's hard to ignore. Evidence on AI and English vocabulary exists, is growing, and largely comes from Latin America, but almost all of it has been gathered from university or high school students in schools with reasonable internet access and devices. There's very little research on basic general education, and even less on Ecuadorian public schools. This wouldn't be so serious if vocabulary weren't precisely the foundation upon which comprehension, writing, and speaking are later built; what's decided pedagogically in a course like seventh grade ends up influencing everything the student does with the language later on. Therefore, more than an academic curiosity, the question of the real effect of these tools in a seventh-grade classroom in a public school is, in reality, an urgent one.

Hence the purpose of this study: to determine the extent to which the use of artificial intelligence tools contributes to the development of English vocabulary in seventh-grade students. Beyond simply adding another case to the already extensive list of studies on AI and language learning, this study aims to fill a specific gap: understanding what happens when these tools are implemented in a classroom that bears no resemblance to the university classrooms where almost all previous research was conducted. To achieve this, the study adopts a mixed-methods approach, combining quantitative measurement of students' lexical progress with a qualitative analysis of their learning experience. This allows the numerical data on vocabulary acquisition to be interpreted alongside the students' and teachers' own perceptions of the learning process.

The article is organized as follows: the Methodology section describes the approach, design, population, and instruments used to collect the information; the Results section presents the findings obtained from both the quantitative and qualitative phases; the Discussion section contrasts these results with the studies reviewed in this introduction; and finally, the Conclusions section summarizes the main contributions of the study and its implications for teaching English vocabulary in public education.

2. MARCO TEÓRICO

It is worth pausing to consider the vocabulary specifically, which is the focus of this study. (Ajisoko, 2020) directly evaluated the use of Duolingo for vocabulary learning and found significant improvements in lexical performance, a finding that directly connects with the tool used in this intervention. Alharbi & Khalil (2023) complement this finding by showing that both students and teachers perceive AI as an effective support for learning vocabulary in English as a second language . Gragera (2024) adds that university students perceive AI as an ally for retaining words and correcting errors, although in contexts with good connectivity, which are hardly comparable to a public school in the region. Guacán et al. (2023) offer an interesting nuance in this regard: AI allows students to be exposed to much more contextualized vocabulary than would fit in a traditional class, an argument that gains even more strength when considering large classrooms with little instruction time, which is precisely the reality of much of Ecuadorian public education.

Other studies point in the same direction from different angles. Ramírez (2025) designed an AI-based intervention and found that being able to adjust the learning pace to each student is one of the most important factors in retaining new vocabulary. Bermúdez et al. (2025) , for their part, insist that AI does not work in isolation: it improves vocabulary and other linguistic skills only when the teacher accompanies it with intentional pedagogical mediation, and not when it is left as an isolated resource that the student uses on their own. Coronado et al. (2025) suggest something similar when studying adaptive AI systems in virtual environments: when the difficulty of the lexical content is adjusted to each student's actual level, frustration decreases and engagement increases—a finding that makes intuitive sense but is rarely measured with data, as they did.

Outside of Latin America, two studies help put this evidence into perspective. Guerrón et al. (2025) show, in a context with much more favorable technological conditions, that chatbots improve both the perceived self-efficacy and the motivation of students of English as a foreign language; this contrast is useful precisely because it reveals how different the starting point can be in scenarios with fewer resources, such as Ecuador. García et al. (2024) , from a broader perspective on the emergence of generative AI in education, warn that its arrival demands new teaching strategies and also a level of digital literacy among teachers that, in many contexts of basic public education, is still far from being achieved.

This international contrast is also maintained when looking at the quantitative evidence on vocabulary itself. The meta-analysis by Hao et al. (2021) , with aggregated data from multiple studies, shows that technological interventions produce positive and consistent effects on lexical performance in students of English as a foreign language. Zhang & Huang (2024) , in an experiment with 52 students divided into experimental and control groups, found that eight weeks with a chatbot was sufficient to improve receptive and productive vocabulary, with sustained gains two weeks later. Yang (2025) , in his review of 30 studies (2015-2023), confirms that this field is growing rapidly but remains fragmented, lacking a shared theoretical framework. Jomaa et al. (2024) found that age and educational level modulate the effect of AI on vocabulary learning, preventing the automatic extrapolation of university results to basic education.

When the evidence focuses on younger students, the international landscape is also limited, but it offers some useful insights. Hsu et al. (2024) found that AI-based image recognition, combined with self-regulation strategies, improves vocabulary acquisition and reduces learning anxiety. Similarly, Chen et al. (2022) integrated speech recognition with dynamic assessment in fifth-grade students in Taiwan and found improvements in oral skills along with reduced anxiety and cognitive load—evidence that supports the use of AI in populations as young as that in this study. Binhomran and Altalhab (2021) tested augmented reality with young learners and reported improvements in vocabulary, although they cautioned that the novelty of the technology could inflate initial enthusiasm. Tai and Chen (2024) found benefits of generative chatbots for oral skills in elementary school students. AlTwijri & Alghizzi (2024) add that AI also affects motivation, engagement, and anxiety, factors that weigh as heavily as instruction itself in a seventh-grade classroom. Liu et al. (2025) , with their mixed-methods design on AI-mediated informal learning in Chinese students, illustrates the type of methodological approach this study adopts, although applied to a university context.

3. METODOLOGÍA

Research approach and design

This study adopts a mixed-methods approach with a sequential explanatory design (DEXPLIS): in the first phase, quantitative data on students' lexical progress are collected and analyzed, and in the second phase, a qualitative perspective is used to delve into the learning experience of these same students and their teacher, with the aim

of interpreting and making sense of what the numbers from the first phase alone cannot explain. The quantitative phase follows a single-group, pre-experimental design, with measurements taken before and after the intervention (pretest-posttest), without a control group, since the study is conducted under the ordinary conditions of an existing classroom and does not seek to isolate variables in a laboratory setting, but rather to describe what happens when these tools are incorporated into a real classroom.

Context and participants

The study was conducted at the "Guayaquil , Emilio Clemente Huerta school" in a seventh-grade section of Basic General Education. All 27 students enrolled in the class, aged 11 to 12, participated. They were selected using census sampling; that is, the entire group was included rather than a smaller sample, given the already small and manageable class size. Students with less than 80% attendance during the intervention period were excluded from the quantitative phase to prevent incomplete data from distorting the pretest-posttest comparison. In addition to the students, the regular English teacher participated in the qualitative phase, providing a perspective that allowed for a comparison between the students' reports and the teacher's daily observations.

Data collection instruments

For the quantitative phase, a 20-item vocabulary test was designed, based on the vocabulary list for seventh grade according to the national English curriculum. The test included multiple-choice and word-image matching items. The same test was administered as a pre-test and post-test to allow for direct comparison. The instrument was reviewed by two English teachers with experience at this level, as a form of content validation through expert judgment, before its implementation. For the qualitative phase, a six-question semi-structured interview guide was developed for the students, focusing on their experience using the AI tool, and a parallel four-question guide was developed for the teacher, focusing on her perception of the changes observed in the classroom.

Procedure

The intervention lasted eight weeks, with two 40-minute sessions per week, integrated into the regular English class schedule. During these sessions, students used Duolingo

to support the practice and review of vocabulary corresponding to the units in the textbook. A vocabulary test was administered before the intervention to establish a baseline; at the end of the eight weeks, the same test was administered as a post-test. In the week following the post-test, interviews were conducted with a subgroup of eight students, intentionally selected to represent different performance levels (high, medium, and low) on the test, along with an interview with the teacher. Informed consent was obtained from the parents or guardians, as was the assent of the students, and authorization was granted by the school authorities for the study.

Data analysis

The quantitative data were analyzed using a paired-samples t-test, comparing the scores obtained in the pretest and posttest, with the support of the SPSS program; the effect size (Cohen's d) was also calculated to quantify the magnitude of the change beyond its statistical significance. The interviews were transcribed and analyzed using thematic analysis, identifying emerging categories related to motivation, difficulties encountered, and the general perception of the tool's use, so that these qualitative findings could be interpreted in dialogue with the quantitative results obtained in the first phase.

4. ANÁLISIS DE RESULTADOS

The results presented below address the central objective of this research: to determine the extent to which the use of artificial intelligence tools contributes to the development of English vocabulary in seventh-grade students in Basic General Education. Following the sequential explanatory mixed-methods design adopted in the study, the analysis is organized into three phases: the quantitative results derived from the pretest-posttest comparison, the qualitative findings obtained from the thematic analysis, and finally, the integration of both phases.

Characterization of the participants

The quantitative phase included 27 seventh-grade students, aged between 11 and 12 years. All cases considered in the analysis met the minimum attendance criterion established for the intervention.

Vocabulary was assessed using a 20-item test, administered both before and after the eight-week educational intervention. The total possible score for each measurement ranged from 0 to 20 points.

Performance in the pretest and posttest

The descriptive analysis showed a clear increase in scores after the intervention. In the pretest, students obtained a mean of 10.30 points (SD = 2.25); in the posttest, this mean rose to 13.22 points (SD = 2.61).

In relative terms, the initial score represented 51.5% of the maximum possible score on the test, while the final score reached 66.1%. This translates to an average increase of 2.93 points between the two measurements, or 14.6 percentage points on the total scale.

Table 1. Descriptive statistics of vocabulary scores

Measurement	n	Average	SD	Minimum	Maximum
Pretest	27	10.3	2.25	6	15
Posttest	27	13.22	2.61	7	18
Difference	27	2.93	1.98	-1	6

Note. SD = standard deviation. The difference was calculated as posttest minus pretest. Source: Own elaboration.

In addition to the increase in the average score, there was an expansion of the upper limit of the scores, which rose from 15 points in the pretest to 18 in the posttest. However, individual performance was not entirely uniform: while improvement was predominant, some students showed smaller changes, two maintained the same score, and one even registered a one-point decrease. This variability highlights that, alongside the overall positive progress of the group, different learning paces existed among the students.

Distribution according to performance level

In order to facilitate the pedagogical interpretation of the scores, three operational levels of performance were established: low, for scores below 12 points; medium, for scores between 12 and 15 points; and high, for scores between 16 and 20 points.

Table 2. Distribution of students according to performance level

Level	Pretest n (%)	Posttest n (%)
Low	19 (70.4%)	5 (18.5%)
Medium	8 (29.6%)	16 (59.3%)
High	0 (0.0%)	6 (22.2%)

Total	27 (100%)	27 (100%)

Source: Own elaboration.

The distribution shows a shift towards higher performance categories. Before the intervention, 19 students (70.4%) were at the low level and none reached the high level. After the intervention, the number of students at the low level decreased to five (18.5%), while 16 (59.3%) moved to the medium level and six (22.2%) reached the high level.

These results complement the increase observed in the group average and show that the change was not concentrated in a few participants, but was accompanied by a substantial modification in the overall distribution of performance levels.

Behavior of vocabulary items

The descriptive analysis of the 20 items allowed for a more detailed examination of how responses behaved before and after the intervention. Overall, an increase in correct responses was observed in the post-test, consistent with the increase in overall scores.

Items that were easier from the outset maintained high success rates in the second assessment, while several of the items with greater initial difficulty showed improvement after eight weeks of work. This suggests that the difference between the pretest and posttest is not explained by isolated changes in one or two items, but rather by a general favorable trend in the assessed lexical performance.

Verification of the assumption of normality

Before performing the inferential comparison, the distribution of differences between the posttest and pretest scores was examined. Given the sample size, the Shapiro-Wilk test was used.

Table 3. Normality test of pretest-posttest differences

Variable	W of Shapiro-Wilk	p
Post-test-pretest difference	0.955	0.284

Source: Own elaboration.

The Shapiro-Wilk test result was not statistically significant ($W = 0.955$; $p = 0.284$). Since the p-value exceeded the threshold of 0.05, there was insufficient evidence to reject the assumption of normality of the differences. This allowed for the use of a related-samples parametric test.

Inferential comparison between the pretest and posttest

Once the assumption of normality was verified, the paired samples Student's t-test was applied to determine if the difference between the scores obtained before and after the intervention was statistically significant.

The mean difference was 2.93 points (SD = 1.98), with a 95% confidence interval between 2.14 and 3.71 points, excluding the value zero.

The results of the t-test confirmed that the difference between the pretest and posttest was statistically significant, t (26) = 7.68; p < 0.001.

Table 4. Pretest-posttest comparison using Student's t-test for related samples

Comparison	Average difference	95% CI	t	gl	p	Cohen's d
Posttest – Pretest	2.93	[2.14, 3.71]	7.68	26	< 0.001	1.48

Note. CI = confidence interval; df = degrees of freedom. Positive values represent higher post-test scores.

Source: Own elaboration.

The confidence interval confirms the positive direction of the observed change, while the significance value supports a statistically consistent difference between the two measurements.

Magnitude of the effect of the intervention

Statistical significance allows us to determine if there are differences between measurements, but it does not, by itself, indicate their magnitude. For this reason, the effect size was calculated using Cohen's d for related measures, yielding a value of d = 1.48.

According to conventional interpretation criteria, this value corresponds to a large effect size. This indicates that the difference recorded between the initial and final measurements was not only statistically significant, but also of considerable magnitude within the analyzed group.

Taken together, the increase in the mean, the shift in performance levels, the confidence interval, the significance of the test, and the effect size form a consistent quantitative pattern that supports the improvement in vocabulary after the intervention.

Results of the qualitative phase

In the second phase of the sequential explanatory design, the experience associated with using the artificial intelligence tool was explored in greater depth. The thematic analysis

was organized around three central categories identified in the study: motivation, difficulties encountered, and general perception of the tool's use. Throughout the interpretive process, vocabulary learning served as a cross-cutting theme, allowing these categories to be linked to the results obtained in the quantitative phase.

Motivation and participation

The first category was related to the motivation generated by incorporating digital activities into English classes. The experience with the tool was characterized by a greater willingness to complete vocabulary exercises, particularly when the activities included immediate feedback, interactive repetition, and progression between exercises.

The dynamic approach to the work made vocabulary practice seem distinct from traditional activities, which focused primarily on memorization and repetition. This motivational component proved relevant to understanding the students' sustained participation throughout the intervention period.

Perception of vocabulary learning

The second trend identified was linked to the perceived progress in recognizing and recalling English words. Repeated exposure to vocabulary, along with the opportunity to correct errors during activities, emerged as a key element in understanding the learning experience.

Frequent practice reinforced previously studied words and maintained continuous contact with the vocabulary. This perception closely corresponds to the quantitative increase observed between pretest and posttest scores.

Feedback and autonomy

Another important element was the ability to obtain immediate feedback on correct and incorrect answers during the activities. This feedback facilitated the identification of incorrect responses and allowed for the repetition of certain exercises when necessary.

This process fostered greater autonomy at certain points in the practice, although it did not eliminate the need for pedagogical support. In this sense, the tool functioned primarily as a supplementary resource for practicing previously assigned content within the subject.

Difficulties during the use of the tool

The positive assessment coexisted with various difficulties. Differences were identified in the pace at which students completed the activities, as well as occasional needs for

guidance in understanding instructions or solving certain tasks.

These differences help to interpret, at least partially, the variability found in the quantitative results. While some students showed significant increases in their scores, others showed more modest progress or remained relatively stable.

Teacher mediation

Pedagogical mediation emerged as a cross-cutting component of the experience. The use of the tool did not replace teacher participation, but rather was incorporated as support for regular English teaching activities.

The guidance provided for tasks, the selection of vocabulary, and the support offered when difficulties arose allowed technology to be integrated with the content covered in class. From this perspective, the observed benefits were not solely related to the availability of technology, but above all to the pedagogical integration of the tool within the teaching and learning process.

Table 5. Summary of the thematic analysis of the qualitative phase

Category	Central finding	Relationship with the quantitative phase
Motivation and participation	Greater willingness to participate in interactive vocabulary activities	It helps to interpret participation during the intervention
Vocabulary learning	Perception of greater recognition and recall of words	This coincides with the increase in post-test scores
Feedback and autonomy	Ability to identify errors and repeat activities	It helps to understand the progressive reinforcement of vocabulary
Difficulties	Differences in learning pace and need for guidance	It allows for the interpretation of the variability of individual gains.
Teacher mediation	The tool functioned as a complement to teaching	Contextualize the observed effect within a pedagogically oriented intervention

Source: Own elaboration.

Integration of quantitative and qualitative results

Integrating the two phases allowed for a broader interpretation of the results. From a quantitative perspective, the mean score increased from 10.30 to 13.22 points, a statistically difference ($t(26) = 7.68$; $p < 0.001$) and a large effect size ($d = 1.48$). Similarly, the percentage of students at the low level decreased from 70.4% to 18.5%, while after the intervention, 22.2% of participants reached the high level.

From a qualitative perspective, motivation, repeated practice, feedback, and perceived progress contributed to a better understanding of the quantitative change. In turn, differences in learning pace and the need for pedagogical support help explain why the gains were not uniform across all participants.

The convergence between both phases shows that the increase in scores occurred in a context marked by a largely favorable experience with AI-supported activities. Therefore, the qualitative results not only serve as additional information but also help explain some of the pedagogical mechanisms linked to the quantitatively identified behavior.

In response to the study's objective, the results show that during the eight weeks of intervention, students experienced a considerable improvement in vocabulary, along with a shift to higher levels of performance. The integration of the findings suggests that frequent practice, immediate feedback, motivation, and teacher mediation were relevant elements in this process. In line with this, the evidence establishes a consistent temporal association between the intervention and the observed improvement, while the assessment of its implications in relation to previous research is addressed in the Discussion section below.

5. DISCUSIÓN

In line with the article's purpose—to determine the extent to which the use of artificial intelligence tools contributes to the development of English vocabulary in seventh-grade students—the results confirm that an improvement in lexical performance was observed after an eight-week intervention supported by Duolingo: the mean score increased from 10.30 to 13.22 points, with a significant difference ($t(26) = 7.68$; $p < 0.001$) and a large effect size ($d = 1.48$). Furthermore, the qualitative phase complemented this trend by identifying motivation, repeated practice, feedback, and teacher mediation as key elements of the learning experience.

Analyzing the study considering the literature used in the Introduction session and focusing on the use of artificial intelligence tools—specifically their potential to enhance English language learning when integrated into a pedagogical strategy—we concur with Ayala & Alvarado (2023). These authors point out that these technologies facilitate personalized instruction and immediate feedback, although they caution that their impact depends on implementation conditions. In this study, the improvement

occurred precisely within sessions planned and integrated into the regular class schedule.

Furthermore, the magnitude of the observed change is like that reported by Zhang & Huang (2024), who found improvements in receptive and productive vocabulary after eight weeks of intervention with a chatbot based on language models. However, there are relevant methodological differences: while those authors used an experimental design with a comparison group, the present study adopted a pre-experimental design with a single group. Therefore, although the increase and size of the effect are favorable evidence, they do not allow us to attribute the change exclusively to the tool, since factors such as regular teaching or cumulative exposure to the language could also have influenced it.

In this vein, the observed shift from low to medium and high levels is consistent with Guacán Tandayamo et al. (2023), who argue that AI resources expand opportunities for exposure to contextualized vocabulary. However, the improvement was not uniform among participants: while some showed significant increases, others presented minor or almost negligible gains. This variability qualifies Ramírez's (2025) assertion regarding adaptation to the learning pace and suggests, in line with Coronado et al. (2025), that technological personalization should be analyzed in conjunction with pedagogical guidance, and not as a sufficient property.

Motivation was one of the most prominent elements in the qualitative phase, in accordance with Guagchinga et al. (2024) and Guerrón et al. (2025), who link the use of AI tools with greater motivation and perceived self-efficacy. Furthermore, the interactive nature of the activities and the immediate feedback appear to sustain student interest more effectively than purely rote memorization strategies, which aligns with AlTwijri & Alghizzi (2024) regarding the link between AI and affective factors such as motivation and engagement. Similarly, the ability to identify and correct errors coincides with the findings of Gragera (2024), although in a university population; its consistency with the present findings suggests that feedback also operates as a relevant mechanism for younger students. The age of the participants is precisely what distinguishes this study from much of the available evidence, which is concentrated at university and high school levels (Berdella & Albornoz, 2025; Torres, 2025). Jomaa et al. (2024) show that variables such as age can modulate the effects of AI-mediated vocabulary learning, so the results with university students should not be generalized to

primary and secondary education. In this regard, Binhomran & Altalhab (2021) found similar improvements in young learners, although they caution that some of the positive response may be due to the novelty effect of the technology, which could also explain, at least in part, the motivation observed in this study. Similarly, Tai & Chen (2024) report benefits of generative chatbots for elementary school students, primarily in oral skills, opening a future avenue of research into the relationship between vocabulary, comprehension, and oral and written production.

One finding considered relevant was the role of teacher mediation: the tool functioned as a complement to, not a substitute for, teaching, in line with Bermúdez et al. (2025). This avoids a technocentric interpretation of the results, since the improvement was not simply due to making an application available to students, but rather to organized and supported sessions within the regular context of the subject. This need for support aligns with García-Peñalvo et al. (2024), who point out that AI in education demands new teaching strategies and greater digital literacy among teachers, especially in contexts, such as Latin America, marked by inequalities in connectivity and technological access, which could widen existing gaps if not properly managed.

From a methodological perspective, the results are situated within a field still fragmented in terms of approaches and evaluation procedures (Yang, 2025), and the use of a mixed-methods design, similar to that of Liu et al. (2025), ensured that the statistical difference was not isolated from the students' lived experience during the intervention.

Finally, the findings must be interpreted considering their limitations: the pre-experimental design without a control group prevents establishing causality; the number of participants was small and corresponded to only one group; and the eight-week intervention allows for the identification of short-term changes, but not their sustainability over time. In summary, the results support the pedagogically planned incorporation of artificial intelligence tools as a complementary, not a substitute, resource for the teaching and learning process of English vocabulary.

6. CONCLUSIÓN

This study aimed to determine the extent to which the use of artificial intelligence tools contributes to the development of English vocabulary in seventh-grade students. The results show that, after eight weeks of intervention with Duolingo support, students' vocabulary improved significantly and with a large effect size (d = 1.48), accompanied

by a notable shift towards higher performance levels.

The qualitative phase revealed that this change depended not only on technological availability, but also on a combination of motivation, repeated practice, immediate feedback, and teacher guidance. The tool functioned as a complementary resource, intentionally integrated into the regular activities of the subject, and not as a substitute for instruction.

The convergence between quantitative and qualitative results supports a favorable association between the intervention and the observed lexical progress, although the single-group pre-experimental design, the small sample size, and the short follow-up period do not allow for definitive causal relationships to be established or whether the gains are maintained over time.

With these limitations in mind, the study provides evidence at a relatively unexplored educational level, distinct from the university contexts where much of the literature on AI and English language learning is concentrated. The findings suggest that incorporating AI tools into basic general

education can be a viable way to strengthen vocabulary, provided it is accompanied by clear pedagogical planning and active teacher mediation.

Looking ahead to future research, it would be valuable to expand the number of participants and institutions, incorporate comparison groups, measure vocabulary retention in the medium term, and explore how variables such as initial English proficiency or access to technology influence the results. Overall, this study suggests that artificial intelligence should not be seen as a replacement for traditional teaching practices, but rather as an ally that, when well integrated, can enrich the English teaching and learning process in public education.

7. BIBLIOGRAFÍA

Abusahyon, A.S.E., Alzyoud, A., Alshorman, O., & Al-Absi, B. (2023). AI-driven technology and chatbots as tools for enhancing English language learning in the context of second language acquisition: A review study. *International Journal of Membrane Science and Technology* , *10* (1), 1209-1223. https://www.academia.edu/download/108026768/1810.pdf

Ajisoko, P. (2020). The Use of Duolingo Apps to Improve English Vocabulary Learning. *International Journal of Emerging Technologies in Learning (iJET)* , *15* (7), 149-155. https://www.learntechlib.org/p/217084/

Alharbi, K., & Khalil, L. (2023). Artificial intelligence (AI) in ESL vocabulary learning: An exploratory study on students and teachers' perspectives. *Migration Letters* , *20* (12), 1030-1045.

AlTwijri, L., & 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). https://www.cell.com/heliyon/fulltext/S2405-8440(24)07084-1

Ayala Pazmiño, M., & Alvarado Lucas, K. (2023). Integración de la Inteligencia Artificial en la Enseñanza del Idioma Inglés en Ecuador: Un Camino para Mejorar los Resultados del Aprendizaje. 593 Editorial Digital CEIT, 8 (Extra 3-1), 679-687. https://dialnet.unirioja.es/servlet/articulo?codigo=9124392

Berdella, MEN, & Albornoz, DRM (2025). Estrategias Didácticas Basadas en Inteligencia Artificial para el Aprendizaje del Inglés: Una Revisión Sistemática de la Producción

Bibliográfica 2019-2023. Ciencia Latina Revista Científica Multidisciplinaria, 9 (1), 4538-

4560. https://doi.org/10.37811/cl_rcm.v9i1.16172

Bermúdez, HDR, Medranda, Á. del RP, Ramírez, FLZ, & Muñoz, JDZ (2025). Integración de la inteligencia artificial en la enseñanza del inglés como lengua extranjera. Scientific Code Research Journal, 6 (E1), 2477-2494. https://doi.org/10.55813/gaea/ccri/v6/nE1/828

Binhomran, K., & Altalhab, S. (2021). The impact of implementing augmented reality to enhance the vocabulary of young EFL learners. *The JALT CALL Journal* , *17* (1), 23-44. http://www.castledown.com/journals/jaltcall/article/view/jaltcall.v17n1.304

Chen, C.-H., Koong, C.-S., & Liao, C. (2022). Influences of Integrating Dynamic Assessment into a Speech Recognition Learning Design to Support Students' English Speaking Skills, Learning Anxiety and Cognitive Load. *Educational Technology & Society* , *25* (1), 1-14.

https://www.jstor.org/stable/48647026

Chicaíza, RM, Camacho Castillo, LA, Ghose, G., Castro Magayanes, IE, & Gallo Fonseca, VT (2023). Aplicaciones de Chat GPT como Inteligencia Artificial para el Aprendizaje del Idioma Inglés: Avances, Desafíos y Perspectivas Futuras. Latam: Revista Latinoamericana de Ciencias Sociales y Humanidades, 4(2), 180. https://dialnet.unirioja.es/servlet/articulo?codigo=9586026

Coronado, JKA, Benalcázar, DVA, Herrera, EJM, Ortiz, JMC, García, EML, & Cabrera, FEG (2025). Personalización del aprendizaje mediante sistemas adaptativos de inteligencia artificial en entornos educativos virtuales. Revista Latinoamericana de Calidad Educativa, 2(2), 69-76. https://doi.org/10.70625/rlce/159

García-Peñalvo, FJ, Llorens-Largo, F., & Vidal, J. (2024). La nueva realidad de la educación ante los avances en inteligencia artificial generativa. RIED-Ibero-American Journal of Distance Education, 27(1), 9-39. https://www.redalyc.org/journal/3314/331475280001/331475280001.pdf

Gragera Retuerto, R. (2024). Percepción de los estudiantes universitarios sobre la efectividad de la inteligencia artificial en el aprendizaje del inglés. European Public & Social Innovation Review [ISSN 2529-9824], vol. 9, (julio de 2024). https://doi.org/10.31637/epsir-2024-401

Guacán Tandayamo, RC, Miguez Haro, RE, Lozada Lozada, RF, Jácome Cobos, DI, & Cruz Gaibor, WA (2023). Inteligencia artificial utilizada como recurso para el aprendizaje. Ciencia Latina: Revista Multidisciplinar, 7 (4), 8263-8277. https://dialnet.unirioja.es/servlet/articulo?codigo=9146852

Guagchinga Chicaiza, NW, Borja Salazar, EC y Pacheco Pruna, EM (2024). Inteligencia artificial en el aprendizaje del idioma inglés para estudiantes de nivel Pre-Intermedio (B1). Serie Científica de la Universidad de Ciencias Informáticas, 17 (11), 45-54. http://scielo.sld.cu/scielo.php?pid=S2306-24952024001100045&script=sci_abstract Guerrón, SAT, Oña, MIU, Calvachi, ASV,

Cabezas, MPG, & Leines, MYS (2025). Use of AI-powered chatbots to support English language learning. *Prospherus* , *2* (3), 772-792. https://doi.org/10.63535/dcy9jn77

Hao, T., Wang, Z., & Ardasheva, Y. (2021). Technology-Assisted Vocabulary Learning for EFL Learners: A Meta-Analysis. *Journal of Research on Educational Effectiveness* , *14* (3), 645-667. https://doi.org/10.1080/19345747.2021.1917028

Hsu, T.-C., Chang, C., & Jen, T.-H. (2024). Artificial Intelligence image recognition using self-regulation learning strategies: Effects on vocabulary acquisition, learning anxiety, and learning behaviors of English language learners. *Interactive Learning Environments* , *32* (6), 3060-3078. https://doi.org/10.1080/10494820.2023.2165508

Jomaa, N., Attamimi, R., & Al Mahri, M. (2024). Utilizing artificial intelligence (AI) in vocabulary learning by EFL Omani students: The effect of age, gender, and level of study. *Forum for Linguistic Studies, 6* (5), 171-186.

Juárez, ASM, Rojas, CNR, Zavaleta, JRS y Saravia, PCC (2024). Uso de ChatGPT y aprendizaje de inglés en estudiantes universitarios. Horizontes Journal of Research in Educational Sciences, 8(33), 834-842. http://www.scielo.org.bo/pdf/hrce/v8n33/a22-834-842.pdf

Liu, G.L., Darvin, R., & Ma, C. (2025). Exploring AI-mediated informal digital learning of English (AI-IDLE): A mixed-method investigation of Chinese EFL learners' AI adoption and experiences. *Computer Assisted Language Learning* , *38* (7), 1632-1660. https://doi.org/10.1080/09588221.2024.2310288

Ramírez Pérez, C. (2025). Uso de la inteligencia artificial para promover el aprendizaje del idioma inglés: Una propuesta de intervención. Ensayos: Revista de la Facultad de Educación de Albacete, 40(1), 100-114. https://dialnet.unirioja.es/servlet/articulo?codigo=10372581

Tai, T.-Y., & Chen, HH-J. (2024). Improving elementary EFL speaking skills with generative AI chatbots: Exploring individual and paired interactions. *Computers & Education* , *220* , 105112. https://doi.org/10.1016/j.compedu.2024.105112

Torres, MV (2025). Contribución de la Inteligencia Artificial o plataforma como medio de aprendizaje del inglés para el desarrollo de habilidades lingüísticas

en estudiantes de ingeniería. Ciencia Latina Multidisciplinary Scientific Journal, 9 (6), 3591-3612. https://doi.org/10.37811/cl_rcm.v9i6.21465

Yang, Y. (2025). AI-supported L2 vocabulary acquisition–a systematic review from 2015 to 2023. *Education and Information Technologies* , *30* (13), 17995-18029. https://doi.org/10.1007/s10639-025-13417-8

Zhang, Z., & Huang, X. (2024). The impact of chatbots based on large language models on second language vocabulary acquisition. *Heliyon* , *10* (3). https://doi.org/10.1016/j.heliyon.2024.e25370

UNEMI
UNIVERSIDAD ESTATAL DE MILAGRO

ISSN: 3103-1315

SCIENTIA Research Club
ADVANCING RESEARCH, SHAPING MINDS.

LETTER OF ACCEPTANCE

21/09/2026

CARTA DE ACEPTACIÓN DE ARTÍCULO CIENTÍFICO

Por la presente se certifica que el artículo titulado "Effectiveness of artificial intelligence tools in the development of English language vocabulary: a study with seventh grade students of Basic General Education"
Un trabajo de investigación de los autores: Doris Fabiola Fuentes Rendón y Jorge Zambrano Pachay, siendo su artículo revisado por Doble Par Ciego y Sistema de Doble Revisión Editorial, antes de ser publicado.

El artículo será publicado en la Revista Científica **Research Club Scientia** ISSN: 3103-1315, en la edición julio – septiembre 2026, Volumen 4, Número 3. Verificable en nuestra plataforma:

https://researchclubscientia.com/index.php/files/issue/archive

Ing. Luci Salas Narvaez, MSc.

EDITORA REVISTA RESEARCH CLUB SCIENTIA

T: +593 96 733 3677
W: www.researchclubscientia.com
M: researchclubscientia@gmail.com

DOI: 10.65415/rcs

UNEMI

UNIVERSIDAD ESTATAL DE MILAGRO

¡Evolución académica!

@UNEMIEcuador