

<https://doi.org/10.69639/arandu.v13i2.2354>

Spaced Repetition to Improve Vocabulary Retention in A2-Level EFL

Repetición espaciada para mejorar la retención de vocabulario en inglés de nivel A2

Gissela Mabel Gorozabel Monge

gmgorozabelm@ube.edu.ec

<https://orcid.org/0009-0007-0214-0268>

Universidad Bolivariana del Ecuador
Ecuador - Durán

Frelly María Molina Pinargote

fmmolinap@ube.edu.ec

<https://orcid.org/0009-0000-2686-4497>

Universidad Bolivariana del Ecuador
Ecuador - Durán

Jacqueline Elizabeth López López

jelopezl@ube.edu.ec

<https://orcid.org/0000-0002-1765-8103>

Universidad Bolivariana del Ecuador
Ecuador - Durán

Josué Reinaldo Bonilla Tenesaca

jrbonilla@ube.edu.ec

<https://orcid.org/0000-0002-6748-2345>

Universidad Bolivariana del Ecuador
Ecuador - Durán

Artículo recibido: 17 mayo 2026- Aceptado para publicación: 28 junio 2026
Conflictos de intereses: Ninguno que declarar.

ABSTRACT

The process of teaching and learning English as a foreign language in Ecuador at the public secondary education level faces difficulties, especially in relation to medium-term retention. The overall objective of the study was to determine the efficiency of spaced repetition in English vocabulary retention in A2-level EFL students at a public educational institution in Portoviejo, Ecuador. The study was conducted using a quantitative approach. The research design was quasi-experimental, pretest–posttest design with a control group. The sampling process was intentional non-probabilistic, for a total of 70 students (35 experimental group and 35 control group). The intervention was carried out with spaced repetition activities applied over six weeks. Two data collection instruments were used: an objective test and a Likert-type questionnaire to assess learning perceptions. Descriptive and inferential statistics were used to analyse the results, using the analysis of covariance (ANCOVA) procedure. When initial performance was adjusted, the results indicated that vocabulary retention improved more clearly in the experimental group, although the effect of spaced repetition was conditioned by students' initial vocabulary level.

After the intervention, greater internal consistency was noted in learning perceptions. Spaced repetition was consolidated as an effective pedagogical method for optimising vocabulary memorisation in A2-level students. Similarly, an improvement was observed when activities were carried out in educational environments with limited instruction time, if differences in students' initial knowledge were considered.

Keywords: spaced repetition, vocabulary retention, English as a foreign language, secondary education, lexical learning

RESUMEN

El proceso de enseñanza y aprendizaje del inglés como lengua extranjera en Ecuador a nivel de educación secundaria pública enfrenta dificultades, especialmente en relación con la retención a mediano plazo. El objetivo general del estudio fue determinar la eficiencia de la repetición espaciada en la retención del vocabulario en inglés en estudiantes de EFL de nivel A2 de una institución educativa pública de Portoviejo, Ecuador. El estudio se desarrolló bajo el enfoque cuantitativo. El diseño de investigación fue cuasi-experimental, pretest–posttest y grupo de control. El proceso de muestreo de tipo no probabilístico intencional, para un total de 70 estudiantes (35 grupo experimental y 35 grupo control). La intervención se desarrolló con actividades de repetición espaciada aplicadas durante seis semanas. Se aplicaron dos instrumentos de recolección de datos: una prueba objetiva y un cuestionario tipo Likert, para evaluar las percepciones de aprendizaje. Se utilizó la estadística descriptiva e inferencial para el análisis de resultados, mediante el procedimiento de análisis de covarianza (ANCOVA). Cuando el desempeño inicial fue ajustado, los resultados indicaron que la retención de vocabulario mejoró con mayor claridad en el grupo experimental, aunque el efecto de la repetición espaciada estuvo condicionado por el nivel inicial de vocabulario de los estudiantes. Después de la intervención, se notó una mayor consistencia interna en las percepciones de aprendizaje. La repetición espaciada se consolidó como un método pedagógico eficaz para optimizar la memorización de vocabulario en los alumnos de nivel A2. De igual manera, se pudo determinar una mejora al desarrollarse actividades en ambientes educativos con tiempo de instrucción limitado, siempre y cuando se considere las diferencias en el conocimiento inicial de los estudiantes.

Palabras clave: repetición espaciada, retención de vocabulario, inglés como lengua extranjera, educación secundaria, aprendizaje léxico

Todo el contenido de la Revista Científica Internacional Arandu UTIC publicado en este sitio está disponible bajo licencia Creative Commons Attribution 4.0 International. 

INTRODUCTION

In education systems with a longer tradition of foreign language teaching, learning English was treated as a gradual process, sensitive to time and memory (Masache-Escobar & Enciso, 2023). In Europe, Asia and North America, vocabulary was no longer understood as an accumulation of isolated words (Namaziandost et al., 2023). It was assumed to be the core that underpins linguistic comprehension and production in real contexts of use (Li, 2025).

International research showed a constant concern for the fragility of lexical learning (Alisoy & Sadigzade, 2025). Teaching words did not guarantee their permanence. Many students passed immediate assessments and, weeks later, were unable to recall the same content (Namaziandost et al., 2020). This phenomenon led to a rethinking of traditional teaching practices focused on immediate repetition and accelerated content coverage (Lee et al., 2021).

In this context, spaced repetition was analysed as a strategy aligned with the principles of human memory functioning (Zhang, 2025). The technique proposed reviewing information at carefully planned, progressive intervals (Alhazmi, 2024). The aim was to reinforce the memory trace before forgetting became consolidated. Experimental studies conducted on different continents reported that this approach allowed for more stable retention and more accurate retrieval of learned vocabulary (Wang & Lee, 2021).

Educational programmes aligned with international language reference frameworks incorporated this strategy with consistent results. Evidence showed that improvement did not depend on increasing the number of words taught. It depended on how time was managed (Li, 2025). This finding generated interest in contexts where instruction time is limited, but the demand for results remains high (Al-Khasawneh, 2023).

In Latin America, teaching English as a foreign language was incorporated as an educational policy in secondary education (Cajas & Cherres, 2024). The region adopted common reference frameworks to define levels of proficiency (Acosta et al., 2022). The objective was clear: to enable students to communicate in English in a basic and functional way.

The reality showed a clear gap between the curriculum's intention and the results obtained. Several regional studies pointed out that vocabulary was one of the main obstacles to linguistic progress. Students learned new words regularly. However, they quickly forgot them. Learning lacked stability (Vargas-Saritama & Celi, 2024).

The prevailing teaching practices explained part of the problem. Instruction focused on the initial presentation of vocabulary. Subsequent reviews were scarce. Time was treated as a neutral variable. This omission had visible consequences on medium-term retention (Parra et al., 2024).

Colombia, Chile, and Brazil began experimenting with distributed revision strategies. These initiatives were based on models developed on other continents. Their application, however, was limited. Most studies were conducted in higher education or in contexts with

specific technological support. Public secondary education continued to show a notable absence of rigorous quantitative evidence (Alvarez et al., 2024).

In Ecuador, English as a foreign language was established as a compulsory subject within the national curriculum (Salgado et al., 2019). Teaching was aligned with international standards of linguistic competence. However, despite this regulatory framework, results in public institutions continue to show limitations in students' English performance (Esquivel-Gómez & Berthely-Barrios, 2025). Vocabulary is one of the most affected components. In Portoviejo, Manabí, this situation is particularly relevant because English teaching is limited to three hours per week. As a result, students have few opportunities to practise and review the language outside the classroom, which restricts systematic reinforcement of lexical content (Tamayo et al., 2023).

The students benefiting from this study took the A2 level of EFL at Unidad Educativa '18 de Octubre'. They were between 13 and 15 years old. They were at a stage of cognitive development that was still unstable. Study habits were not well established. Their contact with English depended almost entirely on what happened in class (Vargas-Saritama & Celi, 2024).

Observation of the context revealed a recurring pattern. Vocabulary was understood in the short term. After a few weeks, it disappeared. The words did not return. Systematic review was not part of the teaching plan. Time was not treated as a teaching ally (Cajas & Cherres, 2024; Parra et al., 2024).

No local experimental studies were identified that analysed this problem using controlled quantitative designs. The absence of empirical data limited informed decision-making. If this situation persisted, the risk was clear. Students would advance to the next level with accumulated lexical gaps that would be difficult to correct at later stages (Alisoy & Sadigzade, 2025).

Based on the picture described above, the following research question was posed:

To what extent did the application of spaced repetition influence English vocabulary retention in A2-level EFL students at Unidad Educativa '18 de Octubre' in Portoviejo, Manabí?

The question was formulated from a quantitative approach. The focus was on estimating the effect of a specific pedagogical strategy on a measurable learning outcome. The overall aim of the study was to determine the efficiency of spaced repetition in English vocabulary retention in A2-level EFL students. The specific objectives were to assess lexical retention before and after the intervention, apply spaced repetition activities adapted to the A2 level, and compare the results between an experimental group and a control group.

Based on the objectives set, the following hypotheses were established:

Alternative hypothesis (H_1): A2 students who received spaced repetition showed significantly higher vocabulary retention, measured by the post-test score adjusted for the pre-test, compared to students who received traditional instruction.

Null hypothesis (H_0): There were no statistically significant differences in vocabulary retention, measured by the post-test score adjusted for the pre-test, between A2 students who received spaced repetition and those who received traditional instruction.

MATERIALS AND METHODS

Research Focus

The study was conducted under a quantitative approach, as its purpose was to measure the effect of spaced repetition on English vocabulary retention through numerical evidence obtained before and after the intervention. This approach was appropriate because the research sought to compare the performance of two groups, estimate changes in vocabulary retention, and test hypotheses using descriptive and inferential statistical procedures (Mendoza et al., 2021). The independent variable was the application of spaced repetition as a pedagogical strategy, while the dependent variable was vocabulary retention in A2-level EFL students. Therefore, the methodological structure was aligned with the need to determine whether systematic review at planned intervals produced measurable differences in students' lexical performance (Bruhn, 2021).

Type and design of research

The research followed a quasi-experimental pretest–post-test design with a non-equivalent control group. This design was selected because the study was conducted in a real school context, where random assignment of students was not feasible. For this reason, the research worked with two intact class groups: one experimental group and one control group (Alisoy & Sadigzade, 2025). Both groups were assessed before and after the intervention in order to compare their vocabulary retention levels over time.

The experimental group received instruction supported by spaced repetition activities, whereas the control group continued with the usual English teaching process without a systematic review schedule. The pretest made it possible to identify the initial vocabulary level of both groups, while the post-test allowed the study to estimate the changes observed after the six-week intervention. This design was appropriate because it enabled the comparison of group performance while statistically controlling for initial differences through covariance analysis (Creswell, 2021).

Population and sample

The study was conducted at Unidad Educativa “18 de Octubre”, located in Portoviejo, Manabí, Ecuador. The institution offers English as a foreign language at the secondary education level, within a curricular structure aligned with the communicative goals established for public education. The target population consisted of A2-level EFL students enrolled in secondary education during the academic period in which the intervention was implemented.

The sample consisted of 70 students aged between 13 and 15 years. A total intentional non-probabilistic sampling procedure was used because the study was carried out with intact classroom groups and random assignment was not feasible in the institutional context (Velazco et al., 2022). Parallel A was assigned as the experimental group and included 35 students, whereas Parallel B was assigned as the control group and included 35 students. Both groups shared similar curricular conditions, school schedules, institutional resources, and English instruction time, which allowed a methodologically acceptable comparison between them (Krippendorff, 2019).

Procedure and intervention

The procedure was organised in three sequential phases to ensure comparable conditions between the experimental and control groups. In the first phase, during Week 1, the objective vocabulary retention test and the Likert-type questionnaire were administered as pre-test instruments to all participants. Both instruments were applied during regular English class hours, under standardised instructions and similar classroom conditions.

In the second phase, from Week 2 to Week 7, the pedagogical intervention was implemented with the experimental group. The intervention consisted of spaced repetition activities focused on the same set of A2-level target vocabulary. Vocabulary review was organised through planned intervals, following an operational sequence of 1, 3, 7, and 14 days, adapted to the school timetable and classroom availability. Each session lasted approximately 40 minutes and was integrated into the regular English class. During the same period, the control group continued with the usual English teaching process, without a systematic spaced repetition schedule.

In the third phase, during Week 8, the same instruments were administered again as post-test measures. The post-test application followed conditions equivalent to those of the pre-test to reduce administration bias. This procedure allowed the study to compare changes between the two groups and to estimate the effect of spaced repetition on vocabulary retention while considering students' initial performance.

Data collection techniques and instruments

Data collection was organised through two complementary techniques aligned with the quantitative nature of the study and the operationalisation of the variables (Hinostroza et al., 2025). The first technique corresponded to pedagogical assessment and was implemented through an objective vocabulary retention test administered as both pre-test and post-test. This instrument measured students' lexical performance before and after the intervention, allowing the study to estimate changes associated with spaced repetition over time. The test focused on the recognition, identification, and contextual use of A2-level English vocabulary through structured items with objective scoring criteria. The scores obtained made it possible to compare group performance quantitatively and to apply inferential statistical analyses.

The second technique corresponded to structured self-reporting and was applied through a Likert-type questionnaire designed to examine students’ perceptions of spaced repetition, vocabulary retention, and learning stability. The questionnaire explored subjective dimensions related to memory reinforcement, repeated exposure to vocabulary, and perceived improvement during the intervention process. Both instruments were administered to the 70 participants under similar classroom conditions and during regular English lessons. The pre-test instruments were applied before the intervention phase, whereas the post-test instruments were administered after the completion of the six-week pedagogical process. This methodological organisation ensured consistency between the variables studied, the instruments applied, and the statistical procedures selected for the analysis (Masache-Escobar & Enciso, 2023).

Instrument 1: Objective vocabulary retention test (pre-test–post-test)

Instrument 1, corresponding to the objective vocabulary retention test in English level A2, was designed according to content validity criteria (see table 1). The purpose was to ensure that the items adequately assessed the proposed construction. The vocabulary was selected in accordance with the content covered during the intervention period and the expected level of linguistic competence for the participating students, in accordance with current curriculum guidelines (Alisoy & Sadigzade, 2025).

The instrument was administered twice: first as a pre-test before the intervention and later as a post-test after the six-week instructional process. In both applications, the scoring system remained the same: each correct answer was assigned 1 point, while incorrect or blank answers received 0 points. The total score ranged from 0 to 10 points, which allowed the results to be interpreted both as raw scores and as percentages of vocabulary retention.

Table 1
Objective English Vocabulary Retention Test (Level A2)

English Vocabulary Retention Test – Level A2							
Date: / /							
Student Instructions							
• Read each question and mark the correct answer. • In the open-ended questions, write the meaning in Spanish.							
Item No.	Item Type	Question / Statement	Option A	Option B	Option C	Option D	Correct Answer
1	Multiple choice	What does borrow mean?	buy	borrow (pedir prestado)	sell	pay	B
2	Multiple choice	What does crowded mean?	empty	full of people (lleno de gente)	quiet	closed	B
3	Multiple choice	What does choose mean?	run	choose (elegir)	forget	lose	B

4	Multiple choice	What does expensive mean?	cheap	expensive (caro)	fast	small	B
5	Sentence completion	I feel ____ because I studied all night.	hungry	tired	early	easy	B
6	Sentence completion	This restaurant is very _____. The food costs a lot.	cheap	crowded	expensive	small	C
7	Sentence completion	She ____ the bus at 7 a.m. every day.	arrives	takes	chooses	borrow	B
8	Sentence completion	Today it is _____. There are many people on the street.	quiet	empty	crowded	late	C
9	Short answer	Write the meaning of neighborhood in Spanish.	—	—	—	—	barrio
10	Short answer	Write the meaning of healthy in Spanish.	—	—	—	—	saludable

Source: Elaborated by Gorozabel Monge, Molina Pinargote, López López, and Bonilla Tenesaca.

Scoring system (for Excel / Jamovi)

Correct answer = 1 point

Incorrect / blank = 0 points

Total, score = 0–10

The items included in the test focused on frequently used words in context, appropriate for level A2, avoiding technical vocabulary or linguistic structures that exceeded the expected competence of the students. This process sought to ensure that the answers reflected lexical knowledge and not difficulties associated with the linguistic complexity of the instrument. The suitability of the items was reviewed by English teachers with experience in secondary education, who assessed the clarity, relevance and correspondence between the learning objectives and the questions asked (Cajas & Cherres, 2024).

Given that the same instrument was used as a pre-test and post-test, the possible effect of repeated testing, associated with the students' familiarity with the items, was considered. To reduce this risk, methodological control measures were adopted. First, the order of presentation of the items was modified in the second application, avoiding an identical sequence that could favour answers from memory. Second, a sufficient time interval was established between the

pretest and posttest, corresponding to the total duration of the intervention, which reduced the probability of immediate recall of the answers (Al-Khasawneh, 2023).

Likewise, the emphasis of the teaching process during the intervention was not on memorising the items in the instrument, but rather on progressive vocabulary learning through spaced repetition activities. This decision was made so that the changes observed in the post-test would respond to lexical learning developed over time and not to mechanical repetition of the test. These methodological considerations allowed us to conclude that Instrument 1 validly assessed vocabulary retention and that the differences observed between the pre-test and post-test were not attributable to familiarity with the test, but rather to the teaching process implemented during the study.

Instrument 2: Likert scale questionnaire

Instrument 2, corresponding to the Likert scale questionnaire, was designed to gather students' perceptions of spaced repetition, vocabulary retention, and the learning process experienced during the intervention (see table 2). Given that the population consisted of A2-level English learners aged between 13 and 15, linguistic comprehension was explicitly considered a central criterion in the administration of the instrument (Teng, 2023).

The questionnaire was administered twice, following the same temporal sequence as the objective vocabulary test. The first application was conducted before the intervention to identify students' initial perceptions of vocabulary learning, memory stability, and review practices. The second application was conducted after the six-week intervention to examine whether students' perceptions changed after the spaced repetition activities. This double application made it possible to compare perceived learning conditions before and after the pedagogical process.

Table 2

Questionnaire with Likert-type scale on spaced repetition and lexical learning

Item No.	Variable	Statement	1	2	3	4	5
RE1	Spaced Repetition	Vocabulary was reviewed at different times during the course.	☐	☐	☐	☐	☐
RE2	Spaced Repetition	Words appeared again in later activities.	☐	☐	☐	☐	☐
RE3	Spaced Repetition	Vocabulary practice was spread across several lessons.	☐	☐	☐	☐	☐
RE4	Spaced Repetition	Time between reviews helped me remember words better.	☐	☐	☐	☐	☐
RE5	Spaced Repetition	Reviewing vocabulary over time reduced forgetting.	☐	☐	☐	☐	☐
RV1	Vocabulary Retention (Perceived)	I remembered the meaning of words learned earlier.	☐	☐	☐	☐	☐
RV2	Vocabulary Retention (Perceived)	I recognized vocabulary studied in previous weeks.	☐	☐	☐	☐	☐

RV3	Vocabulary Retention (Perceived)	I could recall words without reviewing immediately before.	☐	☐	☐	☐	☐
RV4	Vocabulary Retention (Perceived)	Vocabulary stayed in my memory during the study period.	☐	☐	☐	☐	☐
RV5	Vocabulary Retention (Perceived)	I felt my vocabulary knowledge was stable over time.	☐	☐	☐	☐	☐
PA1	Learning Perception	I felt more confident when remembering English words.	☐	☐	☐	☐	☐
PA2	Learning Perception	I noticed less vocabulary loss over time.	☐	☐	☐	☐	☐
PA3	Learning Perception	Vocabulary learning felt more consistent.	☐	☐	☐	☐	☐
PA4	Learning Perception	Spaced review was helpful for learning vocabulary.	☐	☐	☐	☐	☐
PA5	Learning Perception	My overall vocabulary performance improved.	☐	☐	☐	☐	☐

Source: Elaborated by Gorozabel Monge, Molina Pinargote, López López, and Bonilla Tenesaca.

The instrument was administered to all 70 students, both in the experimental group and the control group, maintaining the same structure and number of items at both times. The questionnaire had 5 answer options, from 1 the lowest value to 5 the highest. To ensure that the responses reflected real perceptions and not difficulties in interpretation, the questionnaire was administered in Spanish, using clear and accessible language for the participants. This methodological decision sought to reduce biases associated with limited understanding of the foreign language and to protect the validity of the data collected, especially in self-report scales, where semantic clarity is decisive for the quality of the responses (Cabrera-Solano et al., 2023).

The items were formulated to describe specific classroom experiences, avoiding ambiguous structures or technical vocabulary that could lead to divergent interpretations. Before their final application, the statements were reviewed by English teachers, who assessed their clarity, coherence and suitability for the educational context, thus helping to strengthen the validity of the instrument from a pedagogical perspective (Lee et al., 2021; Li, 2025; Zhang et al., 2022).

Reliability Analysis

The reliability of Instrument 2, corresponding to the questionnaire with a Likert scale, was examined using McDonald's omega coefficient (ω) and Cronbach's alpha coefficient (α). The coefficients were applied independently in the pre-test and post-test stages, after the execution of a pilot test (see table 3). This procedure allowed estimating the internal consistency of the items that made up each scale, considering current psychometric criteria in educational research (Pánek, 2022).

Table 3

Reliability coefficients (McDonald's ω and Cronbach's α) of the pre-test and post-test Likert questionnaire

Instrument	Moment	Scale	Reliability	
			ω	α
Likert	Pretest	RE	0.774	0.779
Likert	Pretest	RV	0.814	0.808
Likert	Pretest	PA	0.736	0.729
Likert	Posttest	RE	0.861	0.859
Likert	Posttest	RV	0.801	0.8
Likert	Posttest	PA	0.874	0.872

Source: Elaborated by Gorozabel Monge, Molina Pinargote, López López, and Bonilla Tenesaca based on data processed in Jamovi 2.7.17.

The reliability analysis was conducted to determine the internal consistency of the Likert-type questionnaire in both applications (Creswell, 2021). McDonald's omega coefficient (ω) was used as the main reliability estimate because it is suitable for scales whose items may not contribute equally to the construct measured (Zou & Xu, 2023). Cronbach's alpha coefficient (α) was also reported as a complementary indicator due to its frequent use in educational research. Both coefficients were calculated separately for the three questionnaire scales: spaced repetition (SR), perceived vocabulary retention (VR), and learning perception (LP).

In the pre-test phase, the reliability coefficients showed acceptable internal consistency across the three questionnaire scales. The spaced repetition scale obtained $\omega = 0.774$ and $\alpha = 0.779$, indicating adequate consistency among its items. The perceived vocabulary retention scale presented $\omega = 0.814$ and $\alpha = 0.808$, which reflected a stronger level of internal consistency. The learning perception scale obtained $\omega = 0.736$ and $\alpha = 0.729$, values that were also acceptable for educational research and supported the use of the scale in the initial measurement.

In the post-test phase, the coefficients remained within acceptable and high reliability ranges. The spaced repetition scale increased to $\omega = 0.861$ and $\alpha = 0.859$, while the perceived vocabulary retention scale obtained $\omega = 0.801$ and $\alpha = 0.800$. The learning perception scale presented the highest internal consistency, with $\omega = 0.874$ and $\alpha = 0.872$. These results indicated that the Likert-type questionnaire maintained a stable internal structure in both applications and supported its use as a complementary instrument for interpreting students' perceptions of vocabulary learning and spaced repetition.

Results analysis technique

Jamovi software version 2.7.17 was used for the quantitative analysis of the data. The information obtained from the objective vocabulary test and the Likert-type questionnaire was first organised in an Excel database and then imported into Jamovi for statistical processing. Descriptive statistics were calculated to summarise the pre-test and post-test scores of both

groups, including means, standard deviations, minimum values, and maximum values. Reliability analysis was performed for the Likert-type questionnaire using McDonald's omega and Cronbach's alpha coefficients. Subsequently, inferential analyses were conducted to compare the initial performance of the control and experimental groups and to estimate the effect of spaced repetition on vocabulary retention. The main inferential procedure was analysis of covariance (ANCOVA), with the post-test vocabulary score as the dependent variable, group as the fixed factor, and pre-test vocabulary score as the covariate.

Ethical safeguards were applied throughout the study. The data were coded to protect the identity of the students and to ensure confidentiality during statistical processing. Participation was handled within the institutional framework of the educational unit, and the information was used only for academic and research purposes. In addition, the study followed general principles of research integrity, including responsible data management, absence of declared conflicts of interest, and respect for the privacy and rights of the participants (COPE, 2024).

Normality analysis

Normality was examined before applying the inferential procedures. The Shapiro–Wilk test was used for the pre-test and post-test vocabulary scores, and its results were complemented with the inspection of distribution plots and Q–Q plots (Samoilenko & Osei, 2021). Descriptive indicators of skewness and kurtosis were also reviewed to support the interpretation of the data distribution. Given the sample size and the use of linear modelling procedures, the decision was not based exclusively on the p-value, but on the joint evaluation of statistical and graphical evidence (Umesh & Dubey, 2022).

Covariance analysis

Before proceeding with the covariance analysis, the statistical assumptions associated with this model were verified to ensure the validity of the inferences made. This control was particularly relevant given the quasi-experimental nature of the study and the use of non-equivalent groups. Furthermore, it is relevant due to the usual situation in educational research carried out in real classroom contexts (Samoilenko & Osei, 2021). First, the relationship between the covariate and the dependent variable was examined. The purpose of this was to confirm a pattern compatible with linearity. Graphical inspection and preliminary analysis indicated an adequate relationship between the pretest and posttest vocabulary retention scores, a necessary condition for the application of ANCOVA (Hadi et al., 2023).

Subsequently, the homogeneity of regression slopes was examined by analysing the interaction between the study group and the covariate. This step made it possible to determine whether the relationship between the pre-test and post-test vocabulary scores differed between the control and experimental groups. The verification of this assumption was relevant because a significant interaction would indicate that the effect of the intervention varied according to students' initial vocabulary level (Creswell, 2021). Similarly, homoscedasticity was checked by

approximating the equality of post-test variances between groups. This analysis was based on tests of equality of variances and inspection of the dispersion of residuals, which allowed us to rule out relevant patterns of heterogeneity (Hernández & Mendoza, 2018).

The distribution of the model residuals was examined to identify severe deviations from the assumption of normality. Graphical evaluation using Q-Q plots and descriptive statistics did not reveal any irregularities that would compromise the robustness of the analysis. The independence of the observations was assumed based on the study design and the individual application of the data collection instruments (Sordi, 2021). Verification of these assumptions supported the use of covariance analysis as an appropriate statistical procedure for estimating the effect of spaced repetition on vocabulary retention, controlling for the students' initial performance.

RESULTS AND DISCUSSION

The initial values showed proximity between groups (see Table 4). The dispersion was similar. The starting point was comparable. This scenario reduced the risk of prior biases. The methodological reading was clear. Subsequent differences could not be attributed to initial advantages. This condition strengthened the adjusted causal inference envisaged in the design (Samoilenko & Osei, 2021).

Table 4

Descriptive statistics of the vocabulary retention pretest by group

Descriptive		
	Group	Pretest
N	Control	35
	Experimental	35
Media	Control	4.74
	Experimental	4.97
Standard deviation	Control	0.741
	Experimental	0.822
Minimum	Control	4
	Experimental	4
Maximum	Control	6
	Experimental	6

Source: Elaborated by Gorozabel Monge, Molina Pinargote, López López, and Bonilla Tenesaca based on data processed in Jamovi 2.7.17.

The descriptive pattern changed clearly. The separation between groups widened. The experimental group showed a higher level (see Table 5). Visual contrast suggested an effect associated with the intervention. This evidence justified the use of an adjusted model to estimate the net impact, considering previous performance (Alisoy & Sadigzade, 2025).

Table 5*Descriptive statistics of the post-test of vocabulary retention by group*

Descriptive		
	Group	Posttest
N	Control	35
	Experimental	35
Media	Control	5.57
	Experimental	8.03
Standard deviation	Control	0.502
	Experimental	0.822
Minimum	Control	5
	Experimental	7
Maximum	Control	6
	Experimental	9

Source: Elaborated by Gorozabel Monge, Molina Pinargote, López López, and Bonilla Tenesaca based on data processed in Jamovi 2.7.17.

The descriptive pattern changed clearly after the intervention. In raw-score terms, the control group increased from 4.74 to 5.57 points, whereas the experimental group increased from 4.97 to 8.03 points (see Table 5). Because the vocabulary test had a maximum score of 10 points, these values also indicated that the control group moved from 47.4% to 55.7%, while the experimental group moved from 49.7% to 80.3%. This descriptive percentage reading does not replace the inferential analysis; rather, it provides a pedagogical interpretation of the magnitude of change observed before applying the adjusted statistical model (Namaziandost et al., 2023).

Table 6*t-Test for independent samples of the vocabulary retention pre-test*

t-Test for Independent Samples								
							Confidence Interval al 95%	
	Statistical	gl	p	Effect Size	Bottom	Upper		
Pretest	T the Student	-1.22	68.0	.226	Cohen's d	-0.292	-0.762	0.180

Note. $H_a: \mu_{\text{Control}} \neq \mu_{\text{Experimental}}$

Source: Elaborated by Gorozabel Monge, Molina Pinargote, López López, and Bonilla Tenesaca based on data processed in Jamovi 2.7.17.

The initial variability was distributed in a comparable manner (see Table 7). The case was considered satisfied. This fact reinforced the stability of the previous contrast. It also reduced the

chance of distortion from uneven spread. The initial comparison was methodologically safeguarded (Acosta et al., 2022).

Table 7
Homogeneity test

Levene test for homogeneity of variances				
	F	gl	gl2	p
Pretest	0.0861	1	68	.770

Note. A low p-value suggests a violation of the equal variance assumption

Source: Elaborated by Gorozabel Monge, Molina Pinargote, López López, and Bonilla Tenesaca based on data processed in Jamovi 2.7.17.

The pretest acted as a predictor of final performance (see Table 8). Statistical control was relevant. Significant interaction indicated a conditioned effect. The intervention did not operate uniformly. The impact varied according to the initial level; therefore, this reading provided a differentiated interpretation of the treatment. The model captured that complexity without simplifying it (Zhang, 2025).

Table 8
Results of the analysis of covariance (ANCOVA) for the vocabulary retention post-test

ANCOVA - Posttest					
	Sum of Squares	gl	Mean Square	F	p
Group	0.115	1	0.1149	2.01	.161
Pretest	24.925	1	24.9246	435.93	<.001
Group * Pretest	1.355	1	1.3550	23.70	<.001
Residuals	3.774	66	0.0572		

Source: Elaborated by Gorozabel Monge, Molina Pinargote, López López, and Bonilla Tenesaca based on data processed in Jamovi 2.7.17.

The homogeneity of variances was not fully maintained (see Table 9). The result called for caution as balanced sample sizes mitigated risk. The methodological literature admitted this condition in applied contexts. The interpretation remained prudent and limited (Lee et al., 2021).

Table 9
Verification of ANCOVA assumptions

Levene test for homogeneity of variances			
F	gl1	gl2	p
9.27	1	68	.003

Source: Elaborated by Gorozabel Monge, Molina Pinargote, López López, and Bonilla Tenesaca based on data processed in Jamovi 2.7.17.

The test indicated statistical deviation (see Table 10). The criterion was not based on the p-value alone. The evaluation considered the robustness of the model. The decision integrated

complementary evidence. This approach was consistent with current practices in educational analysis (Al-Khasawneh, 2023).

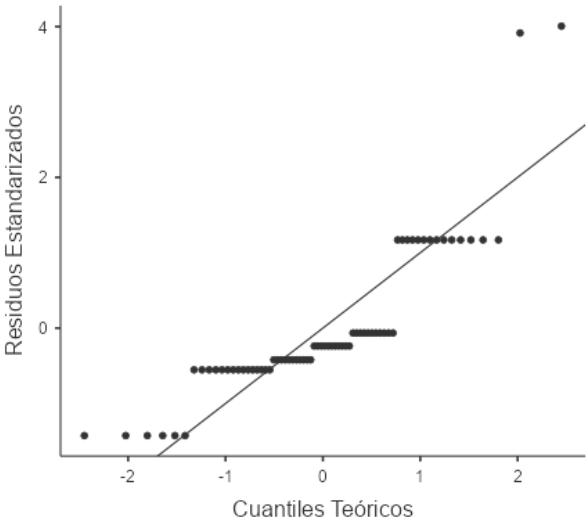
Table 10
Normality Test (Shapiro-Wilk)

Normality Test (Shapiro-Wilk)	
Statistical	p
0.772	<.001

Source: Elaborated by Gorozabel Monge, Molina Pinargote, López López, and Bonilla Tenesaca based on data processed in Jamovi 2.7.17.

Visual inspection of the Q–Q Graph of the standardized residuals of the ANCOVA model showed reasonable alignment (see Figure 1). No critical extremes were observed. The pattern supported the stability of the model. Graphic evidence complemented the formal contrasts. The joint reading made it possible to sustain the operational validity of the adjustment (Esquivel-Gómez & Berthely-Barrios, 2025).

Figure 1
Q–Q graph of the standardized residuals of the ANCOVA model



Source: Elaborated by Gorozabel Monge, Molina Pinargote, López López, and Bonilla Tenesaca based on data processed in Jamovi 2.7.17.

The results of the study allowed us to analyse spaced repetition as a pedagogical strategy conditioned by the initial level of the students. The improvement in vocabulary retention was not uniform. The observed effect depended on the lexical starting point. This pattern coincided with cognitive models that described learning as a process sensitive to repeated memory activation and the prior stability of lexical representations (Teng, 2023). The significant interaction between the group and the pretest revealed that spaced repetition modulated learning differentially. Students did not respond homogeneously to the treatment, as the impact varied depending on the degree of initial vocabulary consolidation. This interpretation was consistent with research indicating that

the effectiveness of distributed review depends on the prior state of the memory trace (Salgado et al., 2019).

From this perspective, spaced repetition did not operate as a neutral application technique; rather, its effectiveness was mediated by the initial conditions of learning. It functioned as an adaptive mechanism. Its effect adjusted to individual differences. This behaviour was described in previous studies in foreign language teaching, where the timing of reviews was more effective when it considered the student's learning trajectory (Alhazmi, 2024).

The improvement in the experimental group also demonstrated immediate repetition as the dominant practice in vocabulary teaching. The literature pointed out that initial exposure, without planned revisions, produces fragile and short-lived learning. In contrast, the temporal distribution of reviews favoured active retrieval and stabilisation of memory, even with limited instructional time (Wang & Lee, 2021). In the Latin American context, these findings took on a particular nuance. Several regional studies highlighted the limited number of hours and scarce exposure to English outside the classroom. Under these conditions, spaced repetition emerged as a viable strategy that did not require expanding content or increasing hours but rather reorganising the use of available time (Masache-Escobar & Enciso, 2023; Namaziandost et al., 2020).

In most cases, students depended on the school environment for their English learning and practice, so the lack of extracurricular practice limited opportunities for spontaneous vocabulary recovery. In other words, the deliberate planning of distributed revisions acted as a compensatory mechanism for natural exposure, offering multiple opportunities for lexical reactivation within the classroom. The results of the Likert-type questionnaire complemented this interpretation. The trends observed in the scales suggested greater internal consistency in students' perceptions after the intervention. The results were consistent with studies indicating that perceptions of learning improve when forgetting is reduced and recall becomes more stable over time (Alisoy & Sadigzade, 2025). The convergence between performance data and perceptions provided methodological triangulation that strengthened the validity of the overall interpretation.

From a methodological standpoint, the significant interaction justified the use of a model with explicit interaction. This approach avoided simplistic interpretations of the treatment effect. Modelling individual variability was consistent with recent recommendations emphasising the need to capture the complexity of learning in applied educational studies (Cajas & Cherres, 2024; Vargas-Saritama & Celi, 2024). Spaced repetition was not configured as a universal solution. It behaved as an effective strategy under specific conditions. Its value lay in time management and repeated memory activation. This reading was aligned with cognitive principles of learning and the structural limitations of educational contexts with restricted resources.

CONCLUSIONS

The study aimed to determine the efficiency of spaced repetition in vocabulary retention among A2-level English learners. The results showed a significant improvement in the performance of the group that received the intervention, once the initial level of lexical knowledge had been controlled for. This evidence provided a direct response to the objective set and confirmed the relevance of the quantitative approach adopted. The descriptive analysis revealed significant and similar data between the two study groups (Teng, 2023). After the intervention was applied, the experimental group achieved a higher level of lexical retention. Therefore, the temporal organisation of revisions has a significant influence on the process of teaching and learning the English language. On the other hand, it can be determined that memory stability is linked to retrieval distributed over time (Alisoy & Sadigzade, 2025; Lee et al., 2021).

The covariance model confirmed that prior performance acted as a predictor of final performance. Furthermore, the significant interaction between the group and the pretest indicated that the efficiency of spaced repetition depended on the student's initial level. The treatment did not produce a uniform effect. This condition provided a more accurate interpretation of the phenomenon and avoided simplistic interpretations of the impact of the strategy applied (Alhazmi, 2024).

The study found significant improvements in vocabulary retention after the intervention to address the objectives. Additionally, the viability of spaced repetition-based activities in an educational context was confirmed. The experimental group's pre-test and post-test results also showed statistically significant differences. As a result, these findings validate the study's internal consistency and the applicability of the tools employed. The null hypothesis, which proposed that there were no differences in vocabulary retention between groups, was rejected considering the empirical findings. After controlling for the initial level assessed in the pre-test, the findings suggested that spaced repetition contributed to higher vocabulary retention under specific initial learning conditions identified through the interaction effect observed in the ANCOVA model. Therefore, support for the alternative hypothesis was interpreted cautiously and in relation to students' prior lexical performance. Given that the participants' initial lexical knowledge conditioned the observed effect, this acceptance was developed cautiously (Esquivel-Gómez & Berthely-Barrios, 2025; Masache-Escobar & Enciso, 2023; Zhang et al., 2022).

Spaced repetition was found to be an effective strategy for improving vocabulary retention in A2-level English learners, provided that individual differences at baseline are considered. This conclusion provided relevant empirical evidence for educational contexts with limited exposure time and established a solid basis for future research aimed at the differentiated adaptation of distributed review strategies.

REFERENCES

- Acosta, H., Cajas, D., & Minchala, E. (2022). Contextualization of Training Input in Multi-Level Replication and Scaling-up Approach in EFL Teacher-Training. *Profile: Issues in Teachers' Professional Development*, 24(2), 31–45. <https://doi.org/10.15446/profile.v24n2.92497>
- Alhazmi, K. (2024). The Effect of Multimedia on Vocabulary Learning and Retention. *World Journal of English Language*, 14(6), 390–399. <https://doi.org/10.5430/wjel.v14n6p390>
- Alisoy, H., & Sadigzade, Z. (2025). Comparative Analysis of Mobile Applications and Traditional Methods in Vocabulary Acquisition. *Educational Process: International Journal*, 17, e2025322. <https://doi.org/10.22521/edupij.2025.17.322>
- Al-Khasawneh, F. M. (2023). The acquisition of foreign language vocabulary: Does spacing effect matter? *The Education and science journal*, 25(3), 174–193. <https://doi.org/10.17853/1994-5639-2023-3-174-193>
- Alvarez, C. L., Tamayo, M. R., & Santos, J. C. dos. (2024). Factors influencing the development of speaking skills among Ecuadorian EFL learners: Teachers' perspectives. *Indonesian Journal of Applied Linguistics*, 14(2), 319–331. <https://doi.org/10.17509/ijal.v14i2.74889>
- Bruhn, K. (2021). A Handbook of Media and Communication Research: Qualitative and Quantitative Methodologies. In *Taylor & Francis Group*. Taylor & Francis Group.
- Cabrera-Solano, P., Ochoa-Cueva, C., & Castillo-Cuesta, L. (2023). Implementing the Engage, Study, Activate Approach Using Technological Tools in Higher Education. *International Journal of Learning, Teaching and Educational Research*, 22(1), 268–282. <https://doi.org/10.26803/ijlter.22.1.15>
- Cajas, D., & Cherres, K. (2024). Challenges and Opportunities of Using Lesson Study in English Language Teaching Among Ecuadorian EFL Teachers. *International Journal of Learning, Teaching and Educational Research*, 23(7), 525–538. <https://doi.org/10.26803/ijlter.23.7.26>
- COPE, C. (2024). *Cooperation between research institutions and journals on research integrity and publication misconduct cases*. Committee on Publication Ethics. <https://publicationethics.org/node/19891>
- Creswell, J. W. (2021). *A Concise Introduction to Mixed Methods Research*.
- Esquivel-Gámez, I., & Berthely-Barrios, J. C. (2025). Training with Multimedia Resources and Verbal Working Memory Tasks: Effects on Vocabulary of English as a Foreign Language. *Mextesol Journal*, 49(4).
- Hadi, M., Martel, C., Huayta, F., Rojas, R., & Arias, J. (2023). Metodología de la investigación: Guía para el proyecto de tesis. In *Metodología de la investigación: Guía para el proyecto*

- de tesis. Instituto Universitario de Innovación Ciencia y Tecnología Inudi Perú.
<https://doi.org/10.35622/inudi.b.073>
- Hernández, R., & Mendoza, C. (2018). *Metodología de la investigación: Las rutas cuantitativa, cualitativa y mixta* (McGraw Hill).
- Hinostroza, E. M. F., Mendoza, D. J., Cejas, M. N., & Trujillo, E. P. P. (2025). Linear regression model to predict the use of artificial intelligence in experimental science students. *International Electronic Journal of Mathematics Education*, 20(1), em0807.
<https://doi.org/10.29333/iejme/15736>
- Krippendorff, K. (2019). Content Analysis: An Introduction to Its Methodology. In *Content Analysis: An Introduction to Its Methodology*. SAGE Publications.
<https://doi.org/10.4135/9781071878781>
- Lee, I.-H., Maechtle, C., & Hu, C.-F. (2021). Enhancing Vocabulary Retention in Low-Achieving EFL Students: Massed or Spaced? *English Teaching and Learning*, 45(4), 415–430.
<https://doi.org/10.1007/s42321-020-00074-y>
- Li, X. (2025). Spaced repetition as a basic structural method for organizing English as a second language teaching. *Porta Linguarum*, 2025(44), 11–28.
<https://doi.org/10.30827/portalin.vi44.30170>
- Masache-Escobar, A., & Enciso, L. (2023). Using Genially to design flashcards to teach English at GBE Elementary. *Iberian Conf. Inf. Syst. Technol., CISTI, 2023-June*.
<https://doi.org/10.23919/CISTI58278.2023.10211743>
- Mendoza, D., Cejas, M., Rivas, G., & Varguillas, C. (2021). Anxiety as a prevailing factor of performance of university mathematics students during the covid-19 pandemic. *Obrazovanie i Nauka*, 23(2), 94–113. Scopus. <https://doi.org/10.17853/1994-5639-2021-2-94-113>
- Namaziandost, E., Mohammed Sawalmeh, M. H., & Izadpanah Soltanabadi, M. (2020). The effects of spaced versus massed distribution instruction on EFL learners' vocabulary recall and retention. *Cogent Education*, 7(1).
<https://doi.org/10.1080/2331186X.2020.1792261>
- Namaziandost, E., Razmi, M. H., Atabekova, A., Shoustikova, T., & Kussanova, B. H. (2023). An Account of Iranian Intermediate EFL Learners' Vocabulary Retention and Recall Through Spaced and Massed Distribution Instructions. *Journal of Education*, 203(2), 275–284. <https://doi.org/10.1177/00220574211031949>
- Pánek, J. (2022). EduChange Methodology. In *EduChange Methodology*. Univerzita Palackého v Olomouci. <https://doi.org/10.5507/prf.22.24461335>
- Parra, O., Reinozo, E., & Granda, M. F. (2024). Exploring Perspectives of Students on Using Mobile Apps to Improve English Vocabulary Through Youtube Videos: A Comparative Study. *International Conference on Evaluation of Novel Approaches to Software*

- Salgado, N., Guaña, J., Escobar, C., & Mohammadfarid, A. (2019). Result of the methodology for learning english distance with the use of TICs. case study: Central university of Ecuador. *Adv. Intell. Sys. Comput.*, 932, 227–233. https://doi.org/10.1007/978-3-030-16187-3_22
- Samoilenko, S., & Osei, K. (2021). *Quantitative Methodologies using Multi-Methods*. Routledge.
- Sordi, J. D. (2021). Design Science Research Methodology. In *Design Science Research Methodology*. Springer International Publishing. <https://doi.org/10.1007/978-3-030-82156-2>
- Tamayo, M. R., Cajas, D., & Sotomayor, D. D. (2023). Using Gamification to Develop Vocabulary and Grammar Among A1 Level of English Students: A Quasi-Experimental Design. *Commun. Comput. Info. Sci.*, 1757 CCIS, 177–190. https://doi.org/10.1007/978-3-031-24978-5_16
- Teng, M. F. (2023). The effectiveness of multimedia input on vocabulary learning and retention. *Innovation in Language Learning and Teaching*, 17(3), 738–754. <https://doi.org/10.1080/17501229.2022.2131791>
- Umesh, B., & Dubey, D. (2022). Research Methodology. In *Research Methodology*. Chapman & Hall. <https://doi.org/10.1201/9781315167138>
- Vargas-Saritama, A., & Celi, V. S. E. (2024). Educaplay as a tool to potentiate English vocabulary retention and learning. *European Public & Social Innovation Review*, 9, 1–16. <https://doi.org/10.31637/epsir-2024-614>
- Velazco, D. J. M., Hinostroza, E. M. F., Moreno, J. E. S., Cerda, J. F. P., & Barros, M. V. S. (2022). Attitudes of Ecuadorian Secondary School Teaching Staff towards Online STEM Development in 2022. *International Journal of Learning, Teaching and Educational Research*, 21(7), 59–81. Scopus. <https://doi.org/10.26803/ijlter.21.7.4>
- Wang, S., & Lee, C. I. (2021). Multimedia Gloss Presentation: Learners' Preference and the Effects on EFL Vocabulary Learning and Reading Comprehension. *Frontiers in Psychology*, 11. <https://doi.org/10.3389/fpsyg.2020.602520>
- Zhang, R., Zou, D., & Xie, H. (2022). Spaced repetition for authentic mobile-assisted word learning: Nature, learner perceptions, and factors leading to positive perceptions. *Computer Assisted Language Learning*, 35(9), 2593–2626. <https://doi.org/10.1080/09588221.2021.1888752>
- Zhang, Y. (2025). Evaluation of the effectiveness of English vocabulary review strategies based on forgetting curves in computer-assisted language learning (CALL) software. *Cogn. Model. Artif. Intell. Conf., AICCONF - Proc. 2025 3rd Cognitive Models and Artificial*

Intelligence Conference, AICCONF 2025 - Proceedings.

<https://doi.org/10.1109/AICCONF64766.2025.11064075>

Zou, P., & Xu, X. (2023). Research Methodology and Strategy. In *Research Methodology and Strategy*. Wiley. <https://doi.org/10.1002/9781394190256>