DOI: https://doi.org/10.23857/dc.v12i1.4713

Ciencias de la Educación
Artículo de Investigación

Instrucción basada en MALL y adquisición de vocabulario AWL: un estudio de intervención con estudiantes de inglés como lengua extranjera (EFL) de nivel B1

MALL-Based Instruction and AWL vocabulary Acquisition: An Intervention Study with B1 EFL Learners

Ensino baseado em MALL e aquisição de vocabulário AWL: um estudo de intervenção com alunos de inglês como língua estrangeira de nível B1

Antus Jose [I]
antus.jose@upec.edu.ec
https://orcid.org/0000-0003-3396-0510

Lenin Renato Roman Tarapues [II]
lenin.roman@ikiam.edu.ec
https://orcid.org/0000-0001-8835-0520

Correspondencia: antus.jose@upec.edu.ec

*__Recibido:__ 23 de enero de 2026 *__Aceptado:__ 15 de febrero de 2026 __* Publicado:__ 05 de marzo de 2026

I. Universidad Politécnica Estatal del Carchi, Ecuador.
II. Universidad Regional Amazónica Ikiam, Ecuador.

Resumen

Esta investigación cuasi-experimental de métodos mixtos determinó la efectividad del paquete Mobile-Assisted Language Learning (MALL) (Quizlet significado/contexto, Wordela ortografía) en mejorar la adquisición de las sublistas 1-2 de la Lista de Palabras Académicas (AWL) entre estudiantes en línea de EFL B1 en una universidad ecuatoriana operada por el estado. Sesenta y cinco estudiantes (experimental n = 34; control n = 31) participaron en una intervención de cuatro semanas en la entrega regular del curso en línea. Los resultados cuantitativos se determinaron mediante pruebas previas y posteriores emparejadas para medir el reconocimiento del significado y la precisión ortográfica. La implementación de la fidelidad se midió a través de registros de la aplicación y registros del LMS. Los hallazgos indicaron que los resultados fueron estadísticamente significativos en el grupo experimental (p = .010) con un tamaño del efecto moderado, mientras que no se indicó ningún progreso significativo en el grupo de control (control). Descriptivamente, las puntuaciones de la prueba posterior fueron mayores en el grupo MALL, pero la diferencia no fue significativa (p = 0,135), lo que indica que el efecto general de las condiciones del estudio fue moderado. Las percepciones generales de usabilidad y apoyo léxico, según lo observado por los estudiantes, fueron generalmente positivas, y los instructores apreciaron el valor pedagógico a pesar de algunos problemas de conectividad y consistencia de uso. En general, los resultados indican que el desarrollo de vocabulario puede estimularse en una situación de inglés como lengua extranjera en línea con la ayuda de la integración estructurada de las herramientas MALL mediante la práctica de la exposición y recuperación repetidas.

Palabras Claves: Aprendizaje de idiomas asistido por dispositivos móviles, adquisición de vocabulario, lista de palabras académicas, estudiantes de inglés como lengua extranjera (EFL), estudio cuasiexperimental, aprendizaje en línea.

Abstract

This quasi-experimental mixed-methods research determined the effectiveness of Mobile-Assisted Language Learning (MALL) package (Quizlet meaning/context, Wordela spelling) in enhancing acquisition of Academic Word List (AWL) Sub-lists 1-2 among online B1 EFL learners at a state-operated Ecuadorian university. Sixty-five students (experimental n = 34; control n = 31) participated in a four-week intervention in regular online course delivery. Quantitative results were determined by matched pre- and post-tests to measure meaning recognition and orthographic accuracy.

Implementation of fidelity was measured through app logs and LMS logs. The findings indicated that the results were statistically significant in the experimental group (p =. 010) with moderate effect size, whereas no meaningful progress was indicated in the control group (control). Descriptively, the post-test scores were higher in the MALL group, but the difference was not significant (p =.135), indicating that the overall effect of the study conditions was moderate. The overall perceptions of usability and vocabulary support as noted by students were generally positive, and the instructors saw pedagogical value despite some issues regarding connectivity and consistency of use. Comprehensively, the results indicate that vocabulary development may be stimulated in an online EFL situation with the help of the structured integration of MALL tools by practicing repeated exposure and retrieval.

Keywords: Mobile-Assisted Language Learning, vocabulary acquisition, Academic Word List, EFL learners, quasi-experimental study, online learning.

Resumo

Esta pesquisa quase-experimental com métodos mistos determinou a eficácia de um pacote de Aprendizagem de Línguas Assistida por Dispositivos Móveis (MALL) (Quizlet para significado/contexto, Wordela para ortografia) no aperfeiçoamento da aquisição das Sublistas 1 e 2 da Lista de Palavras Académicas (AWL) entre alunos de Inglês como Língua Estrangeira (EFL) de nível B1 numa universidade pública equatoriana. Sessenta e cinco alunos (n = 34 no grupo experimental; n = 31 no grupo de controlo) participaram numa intervenção de quatro semanas num curso online regular. Os resultados quantitativos foram obtidos através de pré e pós-testes emparelhados para medir o reconhecimento de significado e a precisão ortográfica. A fidelidade da implementação foi medida através de registos da aplicação e do sistema de gestão de aprendizagem (LMS). Os resultados indicaram que a intervenção foi estatisticamente significativa no grupo experimental (p = 0,010), com um tamanho de efeito moderado, enquanto não se observaram progressos significativos no grupo de controlo. Descritivamente, as pontuações do pós-teste foram mais elevadas no grupo MALL, mas a diferença não foi significativa (p = 0,135), indicando que o efeito global das condições do estudo foi moderado. As perceções gerais de usabilidade e suporte de vocabulário, tal como relatadas pelos alunos, foram geralmente positivas, e os instrutores reconheceram o valor pedagógico, apesar de alguns problemas relacionados com a conectividade e a consistência de utilização. De forma abrangente, os resultados indicam que o desenvolvimento do

vocabulário pode ser estimulado num contexto de ensino de inglês como língua estrangeira online com o auxílio da integração estruturada de ferramentas de aprendizagem assistida por computador (MALL), através da prática repetida de exposição e recuperação.

Palavras-chave: Aprendizagem de línguas assistida por dispositivos móveis, aquisição de vocabulário, lista de palavras académicas, aprendentes de inglês como língua estrangeira, estudo quase-experimental, aprendizagem online.

Introduction

Vocabulary development is crucial for students at the B1 level as they transition from being able to interact with others to handling more linguistic demands. Developing academic vocabulary improves productive skills in language learning. Students who read, write, and argue in higher education should know the words that recur across academic texts.

At Universidad Politécnica Estatal del Carchi in Tulcán, students taking online B1 English language learning courses face various obstacles, such as limited interaction and fluctuating motivation. Most of the time, lessons are delivered via online platform and are often filled with too much information. This type of vocabulary teaching generally involves isolated lists, brief in-class practice, and translation. There is a predictable result: brief gains in familiarity, shallow processing depth, and persistent errors in spelling. Teachers may assign homework, but without mechanisms for spaced retrieval or verified practice, homework is uneven. Students leave courses with partial lexical knowledge that undermines comprehension and written performance in subsequent modules.

Consequently, educators face an evidentiary blind spot: we cannot say reliably whether platforms like Quizlet or Wordela produce meaningful orthographic gains, or whether any observed effects are sustained and attributable to the intervention rather than to extraneous factors. Despite the growing interest in the use of mobile learning tools, there are still gaps in the research on their long-term effects (Chwo et al., 2018). For instance, most studies do not track the retention of the learners after they've completed their studies. Also, there is a lack of evidence on the effectiveness of certain apps, such as Wordela and Quizlet, in Latin American academic institutions.

Through this project, we aim to address these issues by establishing a quasi-experimental intervention study that follows a well-defined cohort of learners. The objective of this study is to evaluate the effectiveness of MALL applications in addressing the local challenges associated with teaching AWL Sub-list 1 and 2. Additionally, it seeks to provide a comprehensive analysis of the effects of these interventions on the learners. Put plainly, the study tests whether mobile-mediated practice produces greater gains than traditional instruction in early AWL acquisition, and whether those gains are equivalent across semantic recognition and orthographic precision.

It continues to be difficult for university-level EFL learners in Ecuador to acquire and retain high-value academic vocabulary, particularly items from the Academic Word List. It is now common practice to teach vocabulary in a decontextualized manner and with limited practice formats, resulting in limited form-meaning associations, contextual recognition, and accurate spelling abilities. Quizlet and Wordela are two mobile tools designed to help learners study vocabulary on the go. They offer spaced practice, multimodal repetition, and performance tracking. These types of applications have not been empirically compared with conventional instruction in Ecuadorian online B1 contexts. Consequently, institutions are unable to integrate MALL tools into their vocabulary curriculum based on evidence. In this study, Quizlet and Wordela are compared with traditional methods of AWL Sub-list 1 and 2 acquisition to fill this gap.

1. What is the baseline level of AWL Sub-list 1 and 2 vocabulary knowledge, including meaning recognition and spelling, among B1-level online EFL students at Universidad Politécnica Estatal del Carchi?
2. To what extent do MALL applications (Quizlet for meaning recognition and context, and Wordela for spelling and orthography) improve students' AWL vocabulary performance compared with traditional instruction?
3. What differences exist in immediate learning outcomes between students using MALL-based instruction and those receiving traditional methods, according to pre and post test results?

To evaluate the effectiveness of mobile-assisted language learning applications, specifically Quizlet and Wordela, in improving Academic Word List Sub-list 1 and 2 vocabulary acquisition, including

meaning recognition and spelling accuracy, among B1 online EFL students at Universidad Politécnica Estatal del Carchi.

- To diagnose the current vocabulary acquisition level of Academic Word List (AWL) Sub-lists 1 and 2 among B1-level pre-intermediate EFL online students at Universidad Politécnica Estatal del Carchi.
- To implement a MALL-based instructional intervention using Quizlet and Wordela, to support the acquisition of the Academic Word List (AWL) Sub-lists 1 and 2 among B1-level pre-intermediate EFL online students at Universidad Politécnica Estatal del Carchi, Tulcán.
- To analyze the effects of MALL-based instructions and traditional vocabulary teaching on students' AWL meaning recognition and spelling outcomes using pre- and post-test results of B1-level pre-intermediate EFL students at Universidad Politécnica Estatal del Carchi, Tulcán.

Literature Review

Vocabulary mastery – especially of academic vocabulary – is crucial for EFL students. The Academic Word List (AWL) by Coxhead (2000) highlights high-frequency academic words that predict academic success. Thus, B1-level students can better comprehend academic texts if they master AWL Sub-lists 1 and 2. It is difficult for readers to acquire AWL words simply by reading them since they are often abstract or formal. To aid vocabulary study, mobile-assisted language learning (MALL) applications like Quizlet (digital flashcards and games) and Wordela (adaptive vocabulary/spelling games) have been proposed. Through the use of gamification features and spaced repetition, MALL provides "anytime, anywhere" practice.

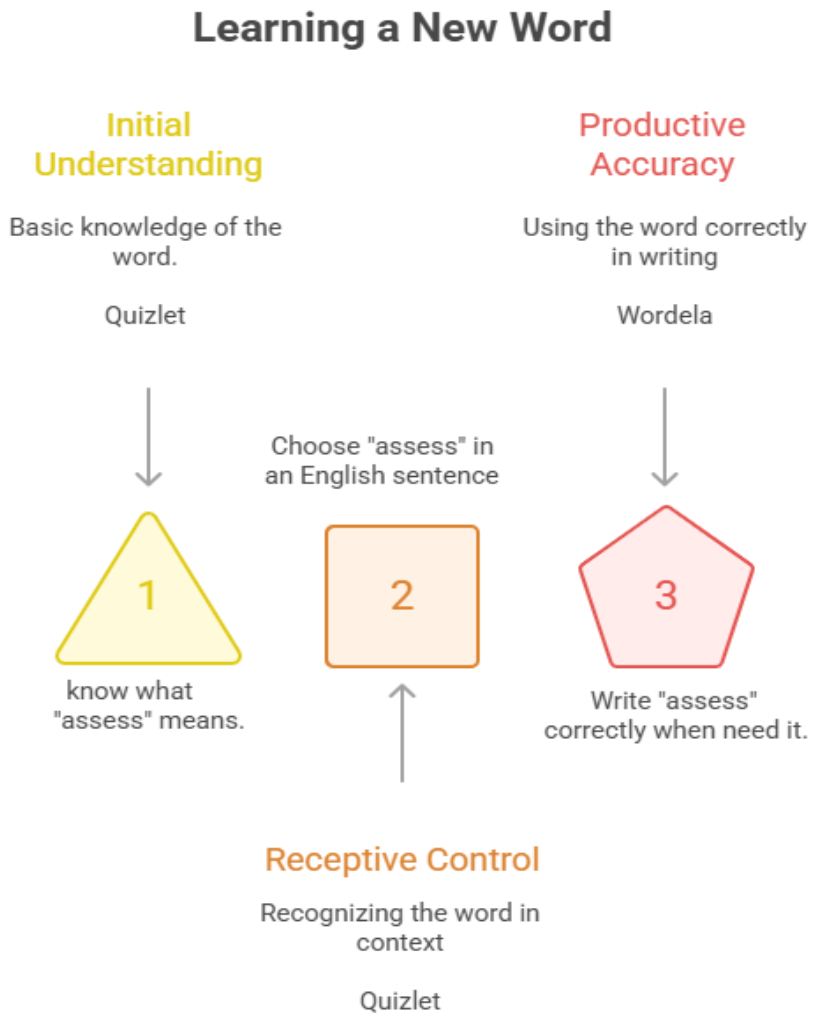
Learners build new vocabulary knowledge by engaging in the active mediation between unknown AWL words and the familiar one, which in this case was facilitated by initial L2-L1 (English-Spanish) meaning mapping in Quizlet (Piaget, 1964). The meaningful learning theory supports the notion that the new lexical items are better retained when anchored to the prior knowledge structures, which is why the short contextual sentences and repeated practices are used to stabilize form-meaning relationships (Ausubel, 1968). In the cognitive approach to technology-based learning, intentionally designed digital tasks that facilitate attention and demand an active processing, e.g. Quizlet

Learn/Test/Match and spelling training in Wordela, can facilitate vocabulary recovery by facilitating manageable and focused practice (Mayer, 2009).

A meta-analysis found that using electronic devices in class significantly increases vocabulary learnin, and recent reviews note that most MALL vocabulary studies report positive gains in word knowledge and learner motivation. Moreover, MALL can foster self-regulated learning and autonomy. Thus, investigating Quizlet and similar apps for AWL acquisition is timely. In particular, this study focuses on the effectiveness of Quizlet (and by extension similar apps like Wordela) for learning AWL sub-lists 1–2, with outcomes of meaning recognition and spelling accuracy, in an Ecuadorian B1 EFL context.

Figure 1.

Progressive Development of Vocabulary Knowledge: From L1 Meaning Recognition to Contextual Use and Spelling Accuracy



Learning a new word is step by step progressive acquisition. Learners first understand the meaning of the word in L1 or any other way, then identify the word in context, and finally can produce the

word with correct spelling. This is the natural order in which vocabulary knowledge emerges (Webb, 2020).

Pellicer-Sánchez's (2020) systematic review substantiates that contextual exposure enhances vocabulary acquisition by illustrating the functional application of words in authentic contexts. Research on spaced repetition suggests that distributed practice is more effective in enhancing long-term retention compared to concentrated study sessions (Nakata & Elgort, 2021). Through a meta-analysis, Uchihara et al. (2019) demonstrated that repeated exposure to target words strengthens form-meaning associations, thereby supporting the use of flashcard-based tools such as Quizlet during the initial stages of learning.

Mobile Apps and AWL Learning: General Findings

Several studies have shown that digital flashcards and MALL tools can significantly improve vocabulary knowledge. For instance, Lukov (2022) found that college EFL learners studying AWL words on Quizlet made significant post-test gains. Okumuş Daedeler (2023) conducted a systematic review indicating that vocabulary studies frequently utilize quizzes and gamified tools to assess learning outcomes. Similarly, Lin & Lin (2019) reached comparable conclusions in their meta-analysis of 33 studies.

Quizlet in EFL Vocabulary Learning

Researchers have compared the effectiveness of digital flashcards with standard cards. They studied whether digital flashcards are better than paper flashcards for vocabulary acquisition. (Dizon & Tang, 2022) Their study showed that users of digital flashcards remembered more vocabulary words than paper flashcard users. When compared to the conventional flashcard method, which also produced positive results, the use of Quizlet offers a more complete and thoroughly constructed learning space. Bakhit et al. (2025) carried out a 10-week quasi-experiment on Malaysian undergraduate students in an EAP course. The apps helped the experimental group to learn the AWL items are Kahoot, Quizlet, and WhatsApp. The post-test scores of the MALL group were much higher than those of the control group. Qualitative interviews also suggested students found the apps engaging. Accordingly, these studies show that the integration of Quizlet and the like is superior or matches regular teaching.

Wordela and Spelling-Focused MALL Tools

Though Quizlet is well-researched, there remains very little scholarly work related to other mobile applications that specifically target spelling in EFL contexts. Wordela is an educational vocabulary and spelling mobile application that utilizes the spaced repetition technique. Gamified tasks are present in the app with an aim to improve the spelling accuracy of learners (eReflect, 2023). There are not any peer-reviewed studies evaluating its effectiveness in academic English as a Foreign Language. Recent evaluations pertaining to mobile lexical acquisition (Okumuş Dağdeler, 2023; Simonnet, 2025) have discerned a plethora of commercial applications; however, none have incorporated Wordela, insinuating that its efficacy remains unexamined within a scientific framework. Empirical studies of developmental spelling also demonstrate that the correct spelling reflects more and more sophisticated orthographic patterns that encode the sound and the meaning as well as that orthographic-morphological knowledge is used to support vocabulary, reading and writing (Colenbrander et al., 2024; Templeton, 2025). For this reason, the present study aims to fill this gap by offering the first empirical evaluation of Wordela and its impact on improving the spelling of Academic Word List vocabulary among B1 EFL learners in an online university setting.

Studies in Latin American Contexts

Empirical MALL research in Latin America is less common, but existing theses and small studies generally support the efficacy of mobile tools. Quizlet was also used as a pre-teaching tool in a rural secondary school by Jaramillo (2021). In her action-research with 36 students, the Quizlet group outperformed on a reading comprehension post-test and students expressed that Quizlet made learning vocabulary “helpful, effective, and dynamic”. These Ecuadorian studies are encouraging but have limitations: There is not clear Ecuador-based empirical evidence combining:

(a) AWL Sublist 1–2 as the target, (b) a MALL intervention, and (c) Quizlet + Wordela together, especially with controlled pre/post outcomes. This research will fill these gaps by focusing on AWL sublists 1 and 2 in an explicit way, by using a control group (traditional instruction) alongside a Quizlet/Wordela group, and by assessing both meaning and spelling.

Methodology

The study adopted a quasi-experimental, mixed-methods design, combining quantitative pre/post measurement with qualitative data to capture process and perception. Two intact B1 online classes

were assigned to experimental and control conditions respectively, with the experimental group receiving a structured mobile-assisted language learning intervention using Quizlet and Wordela and the control group receiving the standard vocabulary instruction. A quasi-experimental approach was selected because random assignment of individual students was impractical within the Universidad Politécnica Estatal del Carchi course structure, yet the research needed comparative, causal inferences about instructional effects.

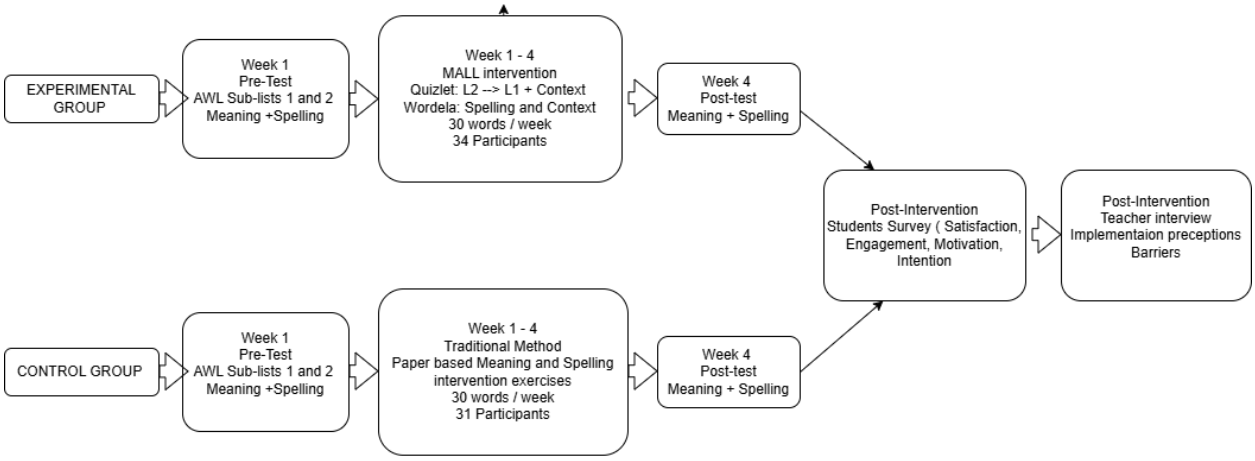
The research was conducted in the university's fully online B1 English programme during a single academic term, a 4-week period from October to December 2025. Instruction and intervention materials were delivered through the university virtual learning environment and students' personal mobile devices, reflecting the ecological conditions of typical course delivery. Quizlet provided activity log where students intervention with the different exercises are being recorded. For wordela, the Moodle platform was used by the students to upload the screen shot of their daily progress.

For Quizlet, two exercises were designed to practice 8 AWL words daily from Monday to Thursday. The first exercise was about meaning recognition from L2 to L1, through practicing the following activities in the Quizlet app in the given order, *digital flash cards, learn, test and matching*(game). The second Quizlet activity is context-based fill-in-the-blank activities and using the right AWL words, which are filled by the student on the *Learn* and *Test* sections.

Afterward, students uploaded the same vocabulary into the Wordela application to practice spelling.

Figure 2.

Study Design and Timeline of the 4-Week Quasi-Experimental Intervention



The participants were drawn from separate sections of the same B1-level course, taught by the same instructor, to minimize instructor effects. To be eligible for inclusion, students must be enrolled in the course and have consented to participating in the study. To achieve reasonable statistical power for medium effects, a target minimum sample of 30 participants per group was planned based on conventional effect-size conventions and prior vocabulary intervention studies (Cohen, 1992; Field, 2013).

Figure 3.

MALL Intervention Workflow: Quizlet Meaning Practice Followed by Wordela Spelling Practice

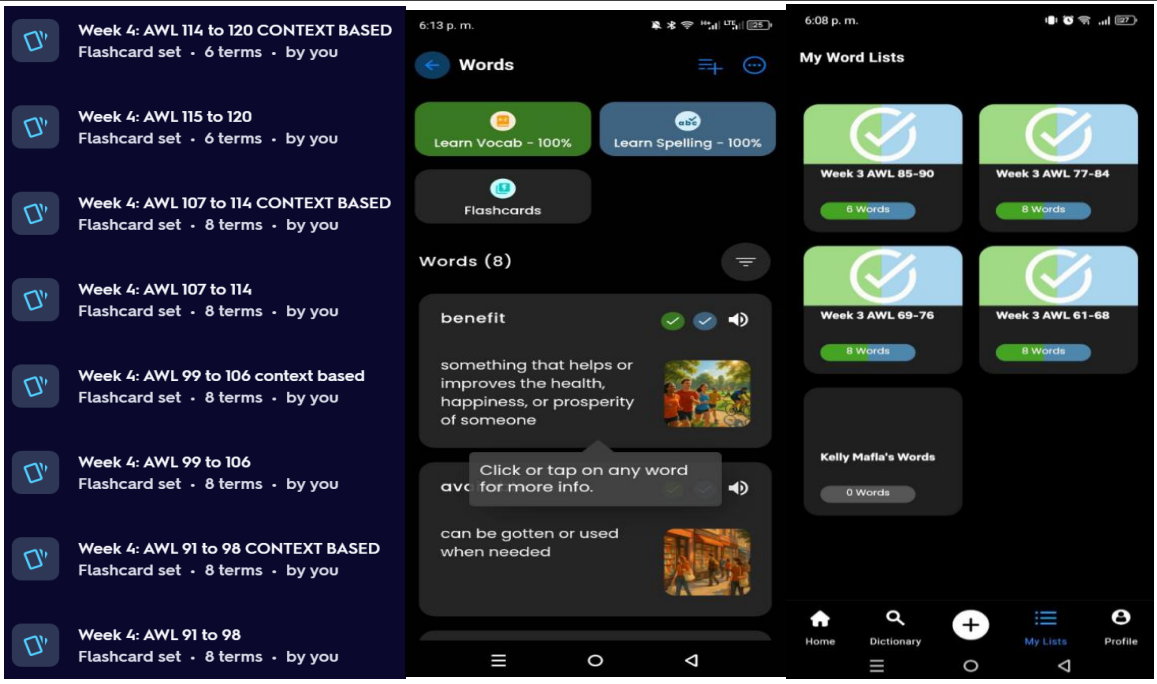


Figure 4.

Quizlet Class Progress Dashboard Used for Fidelity Verification

Fidelity procedures drew on established frameworks for implementation measurement, thereby allowing the study to differentiate exposure from compliance and to link dosage to outcomes (Dane & Schneider, 1998). As part of the control group, the same vocabulary was presented by the teacher, translated, and brief in-class exercises were used. A control arm had worksheets based on paper or a learning management system without algorithmic retrieval spaced over time. Both groups followed a common syllabus and weekly scope to isolate the variable of mobile-mediated practice as the principal difference.

Quantitative data collection comprised a pre-test administered in week 1 and a post-test administered after week 4, with identical test forms counterbalanced across subgroups to reduce practice effects. App log exports, weekly fidelity checklists, and LMS engagement metrics were collected continuously. Qualitative data included a survey to understand about the Satisfaction, Engagement, Motivation across both groups and a short focus interview with the teachers to understand about the methodology they are using to teach vocabulary and the use of MALL applications in their methodology.

An analysis of change scores and adjusted comparisons was performed quantitatively. Paired-samples t tests examined within-group pre/post change, while analysis of covariance models compared post-test scores by condition controlling for pre-test values and relevant covariates such as prior exposure to English and self-reported study time. Cohen's d was reported for within-group changes in ANCOVA, and partial eta squared for ANCOVA results, with assumptions for normality and homogeneity of variance checked (Field, 2013). In sensitivity analyses, relationships between app-documented dosages and outcomes were investigated. Based on a thematic analysis of qualitative transcripts and open responses, Braun & Clarke (2006) identified themes related to engagement, perceived utility, and barriers to use. By triangulating quantitative and qualitative findings, outcomes were interpreted in light of fidelity and learner experience.

Securing ethical approval from the university's research ethics committee was essential. Participants were permitted to withdraw at any point without academic repercussions, provided they had given informed consent. Application log data were anonymized and stored on secure university servers,

with only aggregated usage metrics being reported. The study design sought to minimize harm and maintain academic fairness by granting the control group access to the application resources upon the study's conclusion.

Results

Table 1

Pre-test and Post-test Scores by Group

Participant	**Control Group**		**Experimental Group**	
	Pre-test	Post-test	Pre-test	Post-test
1	8.89	8.11	6.44	5.33
2	4.11	0.67	6.33	7.67
3	9.11	4.67	5.89	5.78
4	9.11	8.56	6.11	7.11
5	6.00	6.67	4.33	5.11
6	8.11	8.89	8.22	8.78
7	4.33	6.89	3.67	7.67
8	8.33	7.44	5.44	5.22
9	4.78	7.22	5.78	5.00
10	4.00	7.11	6.00	6.67
11	4.78	3.22	6.33	9.22
12	6.22	4.33	4.56	6.33
13	8.56	7.89	4.33	6.33
14	7.22	9.00	5.11	5.89
15	6.00	8.78	6.56	5.78
16	9.00	4.89	6.11	5.78
17	6.67	6.11	4.56	9.33

Participant	Control Group		Experimental Group	
18	7.67	8.78	7.56	6.67
19	6.67	5.33	8.67	8.33
20	4.33	6.33	7.56	6.89
21	5.56	7.89	5.67	8.78
22	6.56	6.44	5.89	6.33
23	5.44	6.00	7.33	6.11
24	6.44	5.11	7.56	8.56
25	6.56	9.22	9.33	7.89
26	5.78	4.56	6.11	6.67
27	4.56	7.22	8.56	7.11
28	5.22	5.00	6.78	7.78
29	8.00	9.56	8.00	8.56
30	2.00	2.56	6.89	8.44
31	4.78	6.33	8.00	7.78
32			7.67	7.89
33			4.89	8.22
34			6.00	8.89
			7.33	7.11

Note. Control group: n = 31; Experimental group: n = 35.

Table 2.

Descriptive statistics

Group	n	Pre M (SD)	Post M (SD)	Gain M (SD)
Control	31	6.29 (1.77)	6.52 (2.08)	0.23 (2.02)
Experimental	34	6.44 (1.37)	7.17 (1.27)	0.73 (1.58)

Note. Elaborated by author.

Figure 5.

Mean Pre-test and Post-test Vocabulary Scores by Group (Control vs. Experimental)

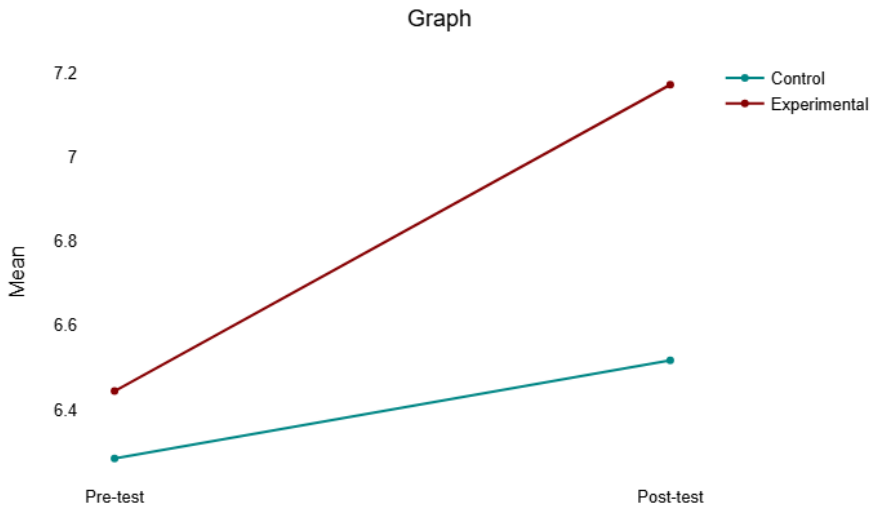


Figure 6.

shows the distribution (median, spread, and outliers) of pre- and post-test scores for both the control and experimental groups.

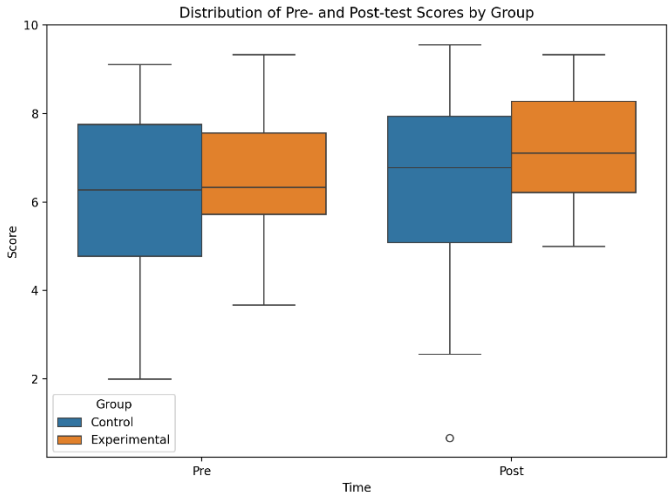
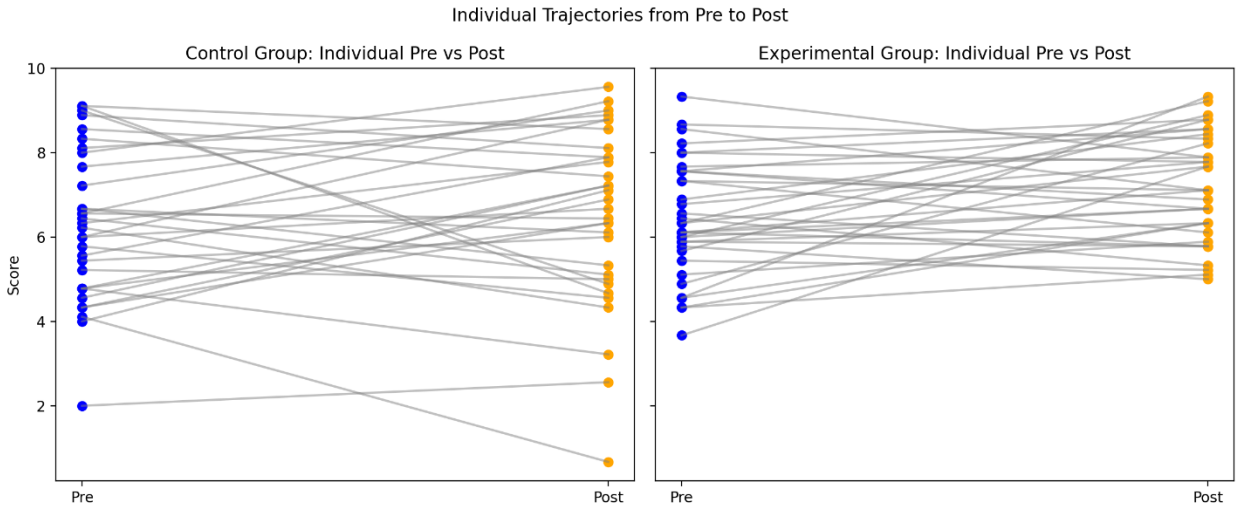


Figure 7

shows each participant's pre- and post-test score, connected by a line, separately for control and experimental groups.



The experimental group more lines tend to go upward (improvement), whereas the control group looks flatter / more mixed.

Paired samples t-test *within each group*

A paired sample t test was conducted to identify control group and experimental group improve over time.

Table 3

	t	*df*	*p*	*Cohen's d*
Pre-test - Post-test	*-2.22*	*66*	*.03*	*0.27*

Note. Elaborated by author.

A paired-samples *t* test indicated a statistically significant increase from the pre-test to the post-test, $t(66) = -2.22$, $p = .03$. This suggests that participants performed better after the intervention. However, the magnitude of the improvement was small (Cohen's $d = 0.27$), meaning the practical impact of the gain was modest.

Independent samples t-test

An Independent samples t-test was performed to understand the MALL instruction really outperform

Figure 8

Mean scores change in pre and post test score for independent samples.

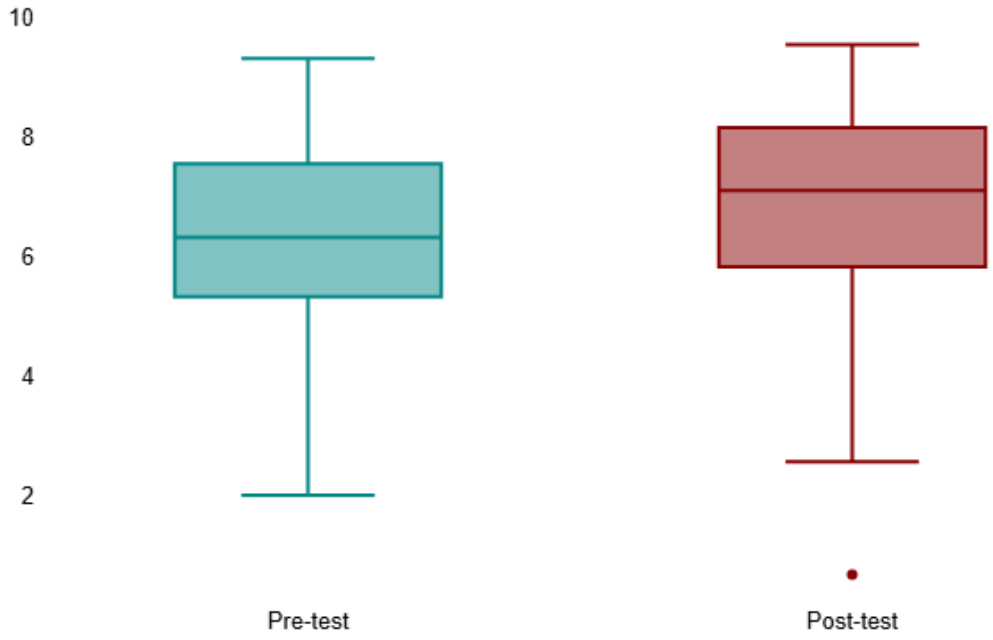


Tabla 4.

	t	df	p	Cohen's d
Equal variances	-1.72	132.00	.087	0.30
Unequal variances	-1.72	130.79	.087	0.30

Note. Elaborated by author.

A two tailed t-test for independent samples (equal variances assumed) showed that the difference between *Pre-test* and *Post-test* with respect to the dependent variable was not statistically significant, $t(132) = -1.72$, $p = .087$, 95.00% confidence interval [-1.05, 0.07]. Thus, the null

hypothesis that there was no difference in the mean value between the two groups was not rejected. The Cohen's d value of 0.30 represents a small effect size.

ANCOVA

Because participants differed in their initial performance, an analysis of covariance (ANCOVA) was conducted on post-test scores using pre-test scores as a covariate to control baseline differences between groups. A one-way ANOVA on post-test scores alone would not account for these initial differences.

Tabla 5

	Sum of Squares	df	Mean Square	F	p
Method	5.71	1	5.71	2.29	.135
Pre-test	30.21	1	30.21	12.14	.001
Error	159.28	64	2.49		

Note. Elaborated by author.

Interpretation

This table shows the results of an ANCOVA. It is tested if *Method* as well as *Pre-test* have an effect on Post-test. You can see each factor's F-value and p-value to see if it explains more variation than you'd expect by chance. Here's a breakdown of each component in the table:

Method

For the factor *Method*, the analysis returned F(1, 64.00) = 2.29 with p = .135. Because the p-value is higher than .05, we conclude that the different Method groups do not differ significantly in their impact on Post-test.

Pre-test

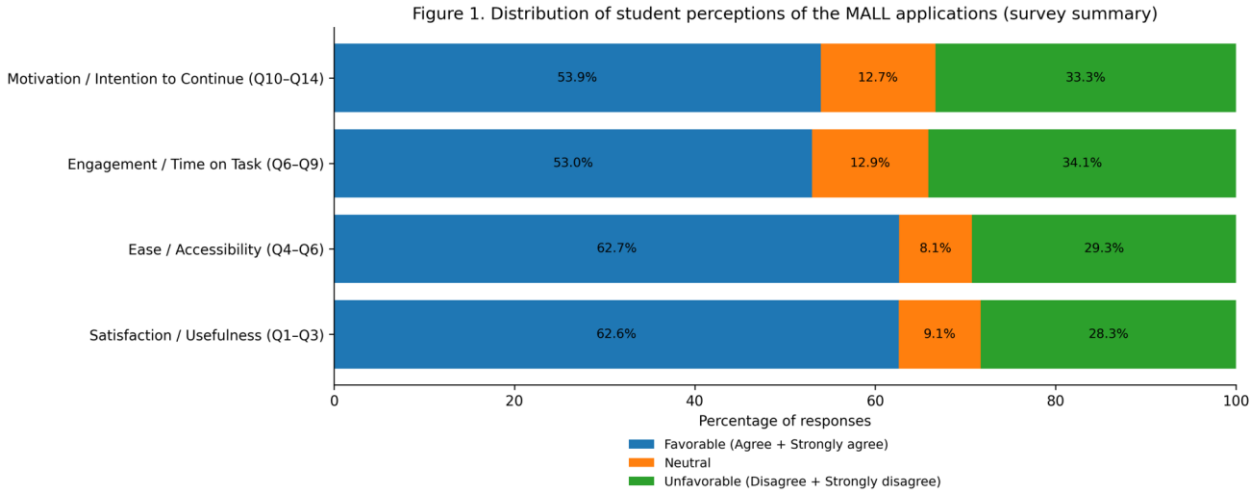
For the covariate *Pre-test*, the analysis returned F(1, 64.00) = 12.14 with p = .001. Because the p-value is lower than .05, we conclude that the variable Pre-test do significantly effect the dependent variable *Post-test*.

An ANCOVA was conducted to compare post-test scores between the control and experimental groups while controlling for pre-test performance. After adjusting for baseline differences, the effect of instructional method on post-test scores was not statistically significant, $F(1, 64) = 2.29$, $p = .135$. In contrast, pre-test scores were a significant covariate, $F(1, 64) = 12.14$, $p = .001$, indicating that initial vocabulary knowledge was a strong predictor of post-test performance.

Students Survey

Figure 9

presents a summary of the post-intervention student survey, which consisted of 14 Likert-scale items (Q1–Q14) designed to capture learners' perceptions of the MALL applications.



Satisfaction / Usefulness (Q1–Q3)

When items were grouped as regards satisfaction and perceived usefulness, then about:

62.6% of responses were favorable (Agree + Strongly agree).

28.3% were unfavorable (Disagree + Strongly disagree).

9.1% were neutral (Not sure / Neither agree nor disagree).

This shows that almost two thirds of students consider MALL application to be useful and satisfactory, though over a quarter are rather dissatisfied. The comparatively minor neutral population is an indicator that majority of the students are certain about using this tool positively or negatively.

Ease / Accessibility (Q4–Q6)

The trend is even more favorable in the case of those items that touch upon ease of use and accessibility:

62.7% of responses were favorable (Agree + Strongly agree).

29.3% were unfavorable (Disagree + Strongly disagree).

8.1% were neutral.

In this case, it is evident that the majority of the students claim that the application is user-friendly and convenient, and the percentage of the Strongly agree position is the biggest of all the constructs. Meanwhile, nearly 1/3 of responses are negative, signifying that a significant subgroup has usability or access obstacles.

Engagement / Time on Task (Q6–Q9)

In the case of engagement and time-on-task items, the distribution of responses was slightly more equal:

53.0% of responses were favorable.

34.1% were unfavorable.

12.9% were neutral.

A little over a half of the students say that they agree or strongly agree that the app encourages engagement and long-term persistence in working on tasks, yet more than one third states the opposite. Engagement is less powerful and polarized than satisfaction and ease, indicating that although most students believe the app to be engaging, a significant portion of them do not perceive it as something that can be very motivating with respect to active or intensive use.

Motivation / Intention to Continue (Q10–Q14)

To be motivated to persist with using the MALL application and to have the wish to do so, the findings are varied:

53.9% of responses were favorable.
33.3% were unfavorable.
12.7% were neutral.

Directly above 50% of the answers show that the students are willing and motivated to keep using the application, but about one third of these students are unwilling to do so. This trend implies that long-term adoption is not as safe as the initial satisfaction or perceived ease of use: although there is a number of students who are positive, there is a certain number that is not planning to implement the app into their current learning activities.

Summary Interpretation

In all four constructs, the total categories of the Agree, + Strongly agree all indicate that the majority of the responses are positive concerning the use of MALL application, particularly the usefulness and ease of use. This is an indication of an overall positive reception and endorses the pedagogical possibilities of the tool.

Teacher Survey Results (Implementation Context and Feasibility)

The teacher survey was carried out on eight English teachers (N = 8). The majority of the respondents were in a public university setting (75.0%), with teaching experience ranging between 0-2 years (12.5%) and over 10 years (37.5). The teacher respondents said that they mostly taught B1 level courses (87.5%), then A1 (62.5%), and A2 (50.0) courses.

The application of Mobile-Assisted Language Learning (MALL) tool, 62.5 percent of teachers indicated they used MALL applications in their English classes and 75.0 percent indicated they used a MALL application in specifically in vocabulary practice (word meaning, spelling or recall). The most used tool was Quizlet (100.0% of teachers), then Memrise (37.5%). Wordela had been reported by 12.5% of teachers. Regarding frequency, teachers showed that students accessed vocabulary apps 1-2 times per month (50.0 percent) or 1-2 times per week (37.5 percent). Specifically, regarding AWL

Sublist 1-2 integration, 12.5% said they used MALL apps to do it, and 50.0% did not use the apps to do it and 25.0% were not sure.

The perceived results were mostly positive: 62.5% of teachers stated that the vocabulary (meaning, spelling, or recall) of students had been improved after using MALL applications, and 25.0% stated that they did not know. Open-ended answers showed that the most frequently mentioned aspects of improvements were better spelling, more correct answers, and more productive use of vocabulary in context (e.g. sentence creation and writing activities). Off-task behaviors (e.g., the use of a translator or AI software) were the most frequent limitations mentioned, followed by connection/technical problems, a lack of time to implement it, and paid subscriptions that prevented using the full features of the apps. Recommended enhancements included more institutionalization of AWL and MALL (e.g. standardized AWL vocabulary lists and curriculum alignment), monitoring/supervision of use, incorporating assessment/grades to enhance compliance, and institutionalizing to minimize translation-based shortcuts.

Discussion

This research explored the hypothesis that an organized Mobile-Assisted Language Learning (MALL) with Quizlet (meaning recognition and contextual practice) and Wordela (spelling practice) could enhance the performance of Academic Word List (AWL) Sub-lists 1-2 in the B1 online EFL learners in comparison to the traditional instructions. This trend of findings was partially in line with the expectation of the study. There was statistically significant increase in pre-test to post-test in the experimental group, $t(34) = 2.73$, $p = .010$, and moderate within-group effect ($dz = 0.46$). On the contrary, no significant change was noted in the control group, $t(31) = 0.65$, $p = .520$. These in-group results can be aligned with the literature that suggests that the vocabulary study based on apps may result in more engagement and provide a chance to practice repeatedly outside the classroom (e.g., Dizon, 2016; Shadiev et al., 2020), and with theoretical propositions that underline the importance of gradual vocabulary acquisition and the benefit of repeating exposure and retrieval during the consolidation of form-meaning associations (Uchihara et al., 2019; Webb, 2020).

Nevertheless, the important quasi experimental question is whether the experimental group improved when compared to control group. Even though the experimental group performed better than the

control group descriptively at the post-test (M = 7.17 vs. 6.52) and had a greater mean gain (0.73 vs. 0.23), conventional statistical significance could not be achieved in the between-group comparisons. The comparison of the gain-scores was found not to be significant, Welch t(58.56) = 1.11, p =.272, and the ANCOVA including the pre-test also did not indicate significant group effect, F(1, 64) =2.29, p=.135(partial e2=.035). This trend indicates that the intervention led to a promising improvement, but the evidence is not enough to assume an evident advantage between groups at =. 05. This finding is not uncommon in practical classroom studies involving intact groups and could be indicative of such limitations as small sample sizes, limited intervention periods, and heterogeneity in independent practice time even where the teaching program is pedagogically sensible.

One of them is that MALL facilitated learning in the same manner in which distributed practice and repeated retrieval lead to a learning process (Cepeda et al., 2006; Roediger and Karpicke, 2006). The intervention plan consisted of conceptually consistent series of vocabulary acquisition models: initial form-meaning access (L2-L1 mapping), contextual recognition, productive accuracy (spelling) which is related to the way the vocabulary knowledge tends to develop over time (Webb, 2020). Moreover, it is commonly known that contextual exposure is useful in consolidating vocabulary and reinforcing functional word knowledge (Pellicer-Sanchez, 2020). However, the limited nature of online learning implementation and practice, in the short period of 4 weeks, probably constrained the extent of between-group separation. In the case that the control condition involved meaningful exposure to vocabulary by teacher-presentation and tasks, one can expect that both groups might have done better, eliminating observable differences. The result of ANCOVA supports the idea that the pre-test results formed a good predictor of the post-test results, which means that the pre-test proficiency level could not be neglected even after the intervention.

The results of the survey of teachers can be used as valuable contextual information to understand the feasibility and implementation requirements. The teacher said they knew about Quizlet well and used AWL Sublist 1-2 with apps relatively poorly, indicating that MALL remains in local practice but has not been structured according to academic vocabulary goals yet. There are also practical barriers, which teachers identified (e.g., connectivity, off-task behavior of students, access limitations) and which tend to dilute dosage of interventions in real environments. These impressions support the rationale of applying fidelity documentation (Quizlet class progress logs; Wordela completion

evidence) to draw the line between exposure and compliance. That is, the mixed-method method of the research study enhances interpretation: where quantitative group effects are not statistically decisive, evidence of implementation makes it clear whether the result is a sign of limited pedagogical effect or dosage and condition limitations.

The lack of statistically significant between-group effect, following the adjustment of pre-test scores, can be, in part, attributed to the differences in how the students were engaged by the MALL activities, since the accomplishment of all the tasks cannot be imposed in the fully online instructional setting.

On the whole, the results justify a tentative conclusion: MALL package seems to help to improve the situation in the experimental class and is estimated as possible and even effective by the instructors, although more significant evidence is required to prove a statistically significant effect in comparison with the traditional instruction on AWL Sublist 1-2 under the existing conditions. Future studies that distinguish between results (meaning vs. spelling), prolong time, and has delayed post-tests can show more definite effects, particularly since the gains in spelling might need time-based practice and feedback loops (Webb, 2020).

Limitations

This paper has a number of limitations that should be taken into account when understanding the results. To begin with, intact groups were employed, and this can include an unmeasured difference even with similarity in the base and the statistical control by ANCOVA. Second, it can be that the four-week intervention was not long enough to measure more significant effects, especially in productive spelling accuracy. Third, despite the intervention being tracked by using application logs and submitted evidence, total adherence to all the MALL activities was not ensured in the fully online setting. Differences between groups could have been dampened by variability in student engagement and time spent in independent study. Lastly, the research was based on post-tests immediately, which did not allow making conclusions about the long-term vocabulary retention.

Curriculum and teaching implications

Regardless of these shortcomings, the trend in the outcomes could have a practical implication on teaching vocabulary online. Consistent practice opportunities, which are hard to achieve based on text-intensive instruction in LMS alone, can be offered by integrating structured MALL activities.

The findings of teacher-surveys also suggest that there is a local necessity of institutional compliance between app-based vocabulary practice and AWL targets. Some of the strategies that curriculum planners can consider are standardizing AWL sublist coverage, integrating app-based practice into the assessment framework to encourage compliance, and simple fidelity monitoring (progress logs/screenshots) to help ensure the quality of implementation in online environments.

Conclusion

The quasi-experimental study examined whether structured MALL instruction (Quizlet meaning recognition and contextual practice and Wordela spelling practice) positively affected the performance of B1 online EFL learners on AWL Sublist 1-2 relative to traditional instruction. The experimental group showed statistically significant pre-post improvement whose within-group effect was moderate, but the control group did not show significant change. Nonetheless, the comparisons between groups (gain-score test and ANCOVA, controlling baseline) were not statistically significant but the results were in favor of the MALL condition.

Based on this, the research finds that the MALL strategy has potential and pedagogical significance in aiding the vocabulary growth in an online university-based course and that more evidence should be gathered to demonstrate that this seems to have a definite benefit in comparison with a traditional course taught under similar exposure. The evidence of the teacher survey indicates that MALL tools are commonly used, and the AWL-centered integration has little presence, which proves the importance of the current study and its focus on the local decision-making process.

Future studies ought to extend the length of intervention, include delayed post-tests in order to gauge retention, and examine meaning and spelling results individually to determine which features of vocabulary knowledge are best supported by each app. Stronger designs (e.g. matching procedures, bigger cohorts, or multi-class replications) and dosage-based analyses can be used to further explain the circumstances under which MALL leads to continually larger gains over traditional vocabulary teaching as well. Overall, the current study offers preliminary local data that structured practice using apps has the potential to enhance the performance of AWL in online B1 classes and suggests practical advice on how evidence-based the introduction of MALL tools in the context of academic vocabulary is planned

References

Chwo, G. S. M., Marek, M. W., & Wu, W.-C. V. (2018). Meta-analysis of MALL research and design. System, 74, 62–72. https://doi.org/10.1016/j.system.2018.02.009

Coxhead, A. (2000). A new academic word list. TESOL Quarterly, 34(2), 213–238. https://doi.org/10.2307/3587951

Piaget, J. (1964). Development and learning. Journal of Research in Science Teaching, 2(3), 176–186. https://doi.org/10.1002/tea.3660020306

Ausubel, D. P. (1968). Educational psychology: A cognitive view. Holt, Rinehart and Winston.

Mayer, R. E. (2009). Multimedia learning (2nd ed.). Cambridge University Press. https://doi.org/10.1017/CBO9780511811678

Webb, S. (2020). Incidental vocabulary learning. In S. Webb (Ed.), The Routledge handbook of vocabulary studies (pp. 225–239). Routledge.

Pellicer-Sánchez, A. (2020). Expanding English vocabulary knowledge through reading: Insights from eye-tracking studies. RELC Journal, 51(1), 134–146. https://doi.org/10.1177/0033688220906904

Nakata, T., & Elgort, I. (2021). Effects of spacing on contextual vocabulary learning: Spacing facilitates the acquisition of explicit, but not tacit, vocabulary knowledge. Second Language Research, 37(2), 233–260. https://doi.org/10.1177/0267658320927764

Uchihara, T., Webb, S., & Yanagisawa, A. (2019). The effects of repetition on incidental vocabulary learning: A meta-analysis of correlational studies. Language Learning, 69(3), 559–599. https://doi.org/10.1111/lang.12343

Lukov, T. S. (2022). Using Quizlet for academic vocabulary acquisition. TESOL Working Paper Series, 20, 34–52. https://www.hpu.edu/research-publications/tesol-working-papers/fall-2022/3-quizlet_lukov.pdf

Okumuş Dağdeler, K. (2023). A systematic review of Mobile-Assisted Vocabulary Learning research. Smart Learning Environments, 10, Article 19. https://doi.org/10.1186/s40561-023-00235-z

Lin, J.-J., & Lin, H. (2019). Mobile-assisted ESL/EFL vocabulary learning: A meta-analysis. Computer Assisted Language Learning, 32(8), 878–919. https://doi.org/10.1080/09588221.2018.1541359

Dizon, G., & Tang, D. (2022). Comparing the efficacy of digital and paper flashcards for second language vocabulary learning. Computer Assisted Language Learning, 35(9), 2207–2227. https://doi.org/10.1080/09588221.2021.1876546

Al Shihri, H. B. S. G., Jahan, N., Jaafar, N. A., & Awang, N. (2025). Examining the effect of the integration of multiple mobile-assisted language learning applications on EFL students' academic vocabulary acquisition: A mixed-methods study. Cogent Education, 12(1), 2473229. https://doi.org/10.1080/2331186X.2025.2473229

eReflect. (2023). Wordela: Vocabulary and spelling training platform [Mobile application software]. eReflect.

Colenbrander, D., von Hagen, A., Kohnen, S., Wegener, S., Ko, K., Beyersmann, E., Behzadnia, A., Parrila, R., & Castles, A. (2024). The effects of morphological instruction on literacy outcomes for children in English-speaking countries: A systematic review and meta-analysis. Educational Psychology Review, 36(4), 119. https://doi.org/10.1007/s10648-024-09953-3

Templeton, S. (2025). The implications and applications of developmental spelling after phonics instruction. Education Sciences, 15(2), 195. https://doi.org/10.3390/educsci15020195

Jaramillo Aguilar, S. P. (2021). The effectiveness of pre-teaching vocabulary through Quizlet to improve reading comprehension in a secondary EFL class: An action research study [Master's thesis, Escuela Superior Politécnica del Litoral].

Cohen, J. (1992). A power primer. Psychological Bulletin, 112(1), 155–159.

Field, A. (2013). Discovering statistics using IBM SPSS statistics (4th ed.). SAGE.

Dane, A. V., & Schneider, B. H. (1998). Program integrity in primary and early secondary prevention: Are implementation effects out of control? Clinical Psychology Review, 18(1), 23–45.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative Research in Psychology, 3(2), 77–101.