Curriculum Research
Volume 6, Issue 2
Jun. 20, 2025

Leveraging artificial intelligence for vocabulary development: Effects on recall and retention in English for specific purposes contexts

Article Type:

Original Research

Authors:

Najmeh Maghsoudi[1]

ORCiD: 0000-0003-3404-6077

Article History:

Received: 2025/03/05

Accepted: 2025/06/12

Published: 2025/06/20

Abstract

English for Specific Purposes (ESP) education has gained prominence as a specialized field designed to address the distinct linguistic needs of learners. Although AI-enhanced instruction has been investigated within general English language learning, its implementation in ESP contexts particularly concerning vocabulary recall and retention has remained insufficiently explored. This study sought to evaluate the impact of AI-supported instruction on vocabulary recall and retention among Iranian ESP learners. Based on non-random convenience sampling, a total of 48 undergraduate students in an ESP course from a university in a southeastern province in Iran participated in the present quasi-experimental study. The participants were non-randomly assigned into two different groups: an experimental group (n=24) receiving AI-powered vocabulary instruction through Diffit, and a control group (n=24) receiving traditional print-based vocabulary instruction. Data were collected during a twelve-week intervention through AI-powered instruction of ESP vocabulary. To address the research questions, a univariate repeated measures analysis of covariate (ANCOVA) was employed. Findings revealed that AI-enhanced instruction significantly improved vocabulary recall and retention among ESP learners. The results offer robust evidence supporting the efficacy of AI-based instructional approaches in enhancing vocabulary learning outcomes within ESP settings.

Key Words: Artificial Intelligence, English for Specific Purposes, Vocabulary Recall, Vocabulary Retention

1. Department of English Language Teaching, Ke.C., Islamic Azad University, Kerman, Iran. Email: maghsoudinajmeh@iau.ac.ir

1. Introduction

The integration of Artificial Intelligence (AI) into educational settings has significantly reshaped traditional pedagogical approaches by enabling more personalized, adaptive, and interactive learning experiences. AI-driven tools, such as Intelligent Tutoring Systems (ITSs) and language learning applications, adapt instructional content to individual learner needs, offering real-time feedback and dynamic adjustments to optimize comprehension and self-regulated learning (Li et al., 2024; Sussmann, 2024). These advancements contribute to the accessibility of flexible, learner-centered environments that promote autonomy and sustained engagement (Hwang et al., 2020; Medina, 2024). Given these developments, a plethora of research has been directed towards the role of AI in language education, investigating its potential to enhance instructional efficacy and learner outcomes (Huang, 2022; Kohnke et al., 2023; Kundu & Parida, 2022; Li et al., 2024; Relmasira et al., 2023)

Noteworthy to mention, the effectiveness of AI in education is contingent upon learners' familiarity with digital tools. Research suggests that students with higher AI literacy demonstrate greater confidence and willingness to integrate AI into their learning processes, whereas those with limited exposure to digital technologies may find the experience overwhelming or ineffective (Hao & Liu, 2022; Kundu & Parida, 2022; Rezaei Ali Kamar et al., 2021). Such disparities underscore the importance of providing adequate training and support to ensure equitable access to AI-enhanced language learning. Furthermore, AI technologies have shown to reduce cognitive load by minimizing extraneous processing demands, allowing learners to focus on higher-order tasks (Mayer, 2021; Sweller, 2019). By delivering instant feedback and automated error correction, AI tools alleviate the cognitive burden associated with self-monitoring, enabling learners to dedicate more cognitive resources to content mastery within a supportive and low-anxiety learning environment (Tuan, 2022).

Unquestionably, English for Specific Purposes (ESP) education has gained prominence as a specialized approach designed to address the distinct linguistic needs of learners, equipping them with the language competencies required in professional and academic contexts (Macia, 2012). Characterized by its goal-oriented, highly relevant, and

practical nature, ESP instruction emphasizes language skill development in alignment with learners' career trajectories (Liu, 2017). A defining feature of ESP is its learner-centered approach that necessitates a thorough understanding of domain-specific language use in students' target occupational settings (Alaqlobi et al., 2024; Chen et al., 2020). Given this focus, ESP practitioners must continuously adapt instructional strategies to ensure alignment with evolving demands in ESP settings.

A critical component of ESP proficiency is vocabulary acquisition, as it underpins learners' ability to function effectively in discipline-specific contexts. Mastery of specialized terminology is essential for the development of language skills (listening, speaking, reading, and writing), as it facilitates comprehension and fluency in both oral and written communication (Pokupec & Njerš, 2014). Vocabulary plays a pivotal role in ESP contexts as it equips learners with the lexical precision required for effective communication in specialized academic or professional domains. Coxhead (2022) also highlighted that targeted vocabulary instruction enhances learners' ability to decode complex texts.

In this regard, Schmidt (2010) argued that vocabulary is central to second language acquisition, necessitating frequent encounters and a combination of explicit instruction and incidental learning. This idea emphasizes that the word knowledge develops over time and in layers of form, meaning, and use. In short, vocabulary is not just a building block, but also the very core of achieving communicative competence in a new language. Similarly, vocabulary is at the heart of ESP, as it enables the users to communicate effectively in specialized fields like medicine, business, industry, or science (Al Zahrani & Chaudhary, 2022). In sum, vocabulary learning in ESP is targeted, practical, and tailored to the learners' actual needs, making it a key driver of communicative competence in specific professional settings (Woodrow, 2017).

Within ESP, vocabulary recall and retention are particularly crucial, as they determine learners' capacity to accurately retrieve and apply domain-specific terminology (Nation, 2013; Schmidt, 2010). Vocabulary recall refers to the ability to actively retrieve learned words when needed, whereas vocabulary retention involves the long-term storage and sustained accessibility of lexical knowledge (Nation, 2013). These processes

are interdependent. While recall ensures immediate communicative competence, retention guarantees the durability of vocabulary knowledge, both of which are indispensable for academic and professional success.

In Iran, where English is taught as a foreign language, ESP courses are increasingly vital due to globalization and the demand for professional communication in fields such as medicine, engineering, tourism, industry, and business etc. (Hayati & Jalilifar, 2009). However, conventional vocabulary instruction often relying on rote memorization and decontextualized word lists fails to meet the specialized needs of ESP learners, resulting in persistent proficiency gaps (Gholami & Khosravi, 2022). This shortcoming is particularly acute in tourism English, a subfield of ESP requiring intercultural communicative competence. Despite its practical orientation, tourism English instruction in Iran lacks authentic contextualization and off-campus practice, leaving learners ill-prepared for real-world interactions (Mostafaei Alaei & Ershadi., 2016).

While AI-enhanced vocabulary instruction has been explored in general English contexts (Huang et al., 2021; Kohnke et al., 2023), its application to ESP particularly for retention-focused outcomes has remained underexplored in Iran. This gap is critical, as retention is a prerequisite for the practical application of ESP vocabulary in professional settings (Stockwell, 2013). The current study attempted to address this gap by investigating the impact of AI-enhanced instruction on vocabulary recall and retention among Iranian ESP learners in tourism, with the following research questions:

1. Does AI-enhanced instruction have any significant effect on Iranian ESP learners' vocabulary recall?
2. Does AI-enhanced instruction have any significant effect on Iranian ESP learners' vocabulary retention?

In line with the research questions, the following hypotheses were formulated:

H_{01}. AI-enhanced instruction has no significant effect on Iranian ESP learners' vocabulary recall.

H_{02}. AI-enhanced instruction has no significant effect on Iranian ESP learners' vocabulary retention.

2. Review of the Related Literature

Empirical studies consistently report positive outcomes for AI-assisted vocabulary learning (Kohnke & Moorhouse, 2021; Pham et al., 2024; Silitonga et al., 2024; Sun & Wang, 2020; Valencia et al., 2020; Zhang & Zou, 2020). AI-driven platforms employ adaptive learning algorithms, spaced repetition, and natural language processing (NLP) to tailor vocabulary instruction to individual learners, thereby maximizing engagement and learning effectiveness (Selvi & Thirumoorthi, 2024). These technologies address common challenges in vocabulary acquisition such as lack of contextualization, difficulties in pronunciation, and limited motivation by providing immediate feedback, contextual examples, and gamified elements that sustain learner interest (Cui, 2024).

As in other areas of education, AI is reshaping ESP vocabulary instruction by offering individualized, real-time, and contextually relevant practice opportunities. In ESP contexts, AI applications can be implemented to generate examples, practice tasks, interactive games or quizzes, and even real-world communication scenarios relevant to learners' careers. AI also enables learners to review and expand specialized vocabulary outside the classroom, supports self-study, and helps teachers quickly prepare targeted materials (Coxhead, 2022).

Research has revealed that mastering technical, semi-technical, and general vocabulary is essential for ESP students to comprehend and produce discipline-specific texts (Coxhead, 2022; Pokupec & Njerš, 2014; Woodrow, 2017). Additionally, needs analyses reveal that vocabulary gaps directly hinder occupational competence, underscoring the necessity of context-driven lexical instruction (Dudley-Evans & St John, 1998). ESP pedagogy thus prioritizes sociocultural competence; ensuring learners not only acquire terms but also use them appropriately within professional discourse communities. Ultimately, vocabulary mastery in ESP bridges linguistic and disciplinary knowledge, fostering both academic success and career readiness (Makhmudova & Mashrapova, 2024).

The consensus among scholars (Dudley-Evans & St John, 1998; Makhmudova & Mashrapova, 2024) is that ESP learners require not just technical terminology but also sociocultural competence to use language appropriately in professional settings. This

aligns with broader ESP pedagogy, which emphasizes context-driven instruction. However, the literature does not sufficiently address how AI can facilitate sociocultural learning, as most studies focus on retention and recall rather than pragmatic or discourse-based competence.

The present study is anchored in Cognitive Load Theory (CLT), which posits that learning efficiency is maximized when instructional design aligns with the brain's cognitive architecture, minimizing extraneous load while fostering relevant cognitive processing (Sweller, 1988). AI-driven tools, particularly adaptive learning systems, operationalize CLT principles by tailoring vocabulary instruction to individual proficiency levels, thereby optimizing cognitive load management (Sweller, 2019; Sweller et al., 2019). The application of CLT provides a robust framework for understanding the benefits of AI in vocabulary learning. Adaptive capabilities of AI align well with CLT by minimizing extraneous cognitive load and optimizing schema acquisition. However, empirical validation remains limited—most cited studies (e.g., Kohnke & Moorhouse, 2021; Zhang & Zou, 2020) focus on short-term outcomes rather than long-term cognitive efficiency.

Research on the efficacy of AI-enhanced instruction in improving vocabulary recall and retention has yielded mixed yet promising results. Several studies highlight AI's dual benefits, noting that retention effects tend to be more pronounced (Huang & Zou, 2021; Lin & Huang, 2021; Lin & Vuono, 2019; Lu & Li, 2020; Zou & Xie, 2019). Conversely, other investigations suggest that AI-driven adaptive learning systems enhance both immediate recall and long-term retention by personalizing learning trajectories (Chen & Zhang, 2019; Kohnke & Moorhouse, 2021; Sun & Wang, 2020; Wang & Vásquez, 2021; Zhang & Zou, 2020). For instance, AI platforms facilitate self-paced vocabulary engagement that strengthens recall (Chen & Zhang, 2019), while adaptive tools reinforce retention through immediate corrective feedback and systematic practice (Kohnke & Moorhouse, 2021).

Direct engagement through AI-driven instruction has also contributed to positive ESP vocabulary learning. Studies indicate that learners using AI chatbots outperform peers in traditional settings, both in recall and retention, as these tools enable repeated exposure to vocabulary in meaningful contexts, adapt difficulty levels to learner needs,

and provide instant feedback (Lyu et al., 2024). This active involvement boosts learner motivation and self-directed learning, making vocabulary practice both accessible and engaging—particularly valuable for adult learners and those with professional commitments (Pham et al., 2024).

Empirical evidence further supports these claims. For instance, Silitonga et al. (2024) observed significant improvements in ESP vocabulary acquisition among students using an AI chatbot (Dialogflow), while Valencia et al. (2020) reported enhanced retention and motivation among foreign language learners utilizing multimodal strategies of Memrise. Similarly, Pham et al. (2024) reported that students' use of POE as an AI tool exhibited sustained engagement and perceived its utility favorably, with all participants expressing intent to continue its use.

Surprisingly, the results of some studies show that AI does not always lead to English vocabulary development. For example, Bastani et al. (2024) found that high school students who used AI platforms actually performed worse on vocabulary assessments compared to the participants with no AI access, possibly due to over-reliance on AI. Similarly, Machin-Mastromatteo (2023) suggested that while AI can help, results are mixed, with some learners showing minimal or no improvement in vocabulary when using AI-based approaches over traditional instruction. As another study, Trabelsi (2025) highlighted similar findings, noting over-dependence on AI tools can sometimes harm language learning outcomes, with students showing weaker retention and recall after practicing with AI compared to traditional methods.

While AI offers dynamic, customizable learning experiences (Hwang et al., 2020; Coxhead, 2022), its real-world implementation faces challenges such as over-reliance on technology. Although Silitonga et al. (2024) reported high learner engagement with AI tools, it is unclear whether this translates to autonomous language use in professional settings. Despite the potentials of AI in vocabulary development, variability in outcomes persists. The meta-analysis on AI-assisted language learning conducted by Huang et al. (2022) found that while AI tools consistently enhance retention, their effects on recall and retention depend on design factors like interactivity and personalization.

Regrettably, longitudinal data remain scarce as Alhusaiyan (2025) cautioned; most

studies (Pham et al., 2024; Silitonga et al., 2024) measure outcomes over weeks rather than months, leaving open questions about durability. Empirical validation of AI's long-term cognitive benefits is limited as most studies (Kohnke & Moorhouse, 2021; Zhang & Zou, 2020) focused on immediate gains rather than sustained lexical automation. Such studies presented an optimistic view of the role of AI in vocabulary acquisition. For instance, the studies by Huang & Zou (2021) and Lin & Vuono (2019) highlighted the advantages of AI; however, the mixed results suggest variability in implementation—some AI tools such as Memrise may excel in retention while other AI tools like POE enhance engagement.

3. Method

3.1. Design

The current study adopted a quasi-experimental pretest-posttest control group design. The artificial intelligence instruction was the independent variable and ESP vocabulary recall and retention functioned as the dependent variables. This design was selected as the study sought to investigate the effectiveness of AI-enhanced instructional intervention on vocabulary recall and retention among ESP learners.

3.2. Participants

Using non-random convenience sampling, the study adopted a quasi-experimental design with 57 undergraduate Tourism students (37.2% male, 62.8% female) enrolled in an English for Specific Purposes (ESP) course at a university in a southeastern province in Iran. Participants, aged 19-26 (M= 20.4, SD= 2.1), were native Persian speakers. For Iranian undergraduate learners of Tourism, ESP is a compulsory course administered in the second year of their program.

Prior to the treatment, the researcher administered Oxford Placement Test (OPT) to ensure the homogeneity of the participants. The results of OPT revealed that 48 participants were at the pre-intermediate level of English proficiency, 4 learners were at intermediate level, and 5 participants showed elementary level of English proficiency. As a result of excluding the intermediate and elementary students from the final analysis, the

researcher was left with a sample of 48 pre-intermediate participants who were randomly assigned to an experimental group (n=24) and a control group (n=24). It should be noted that the participants in experimental group were required to avoid exposure to AI tools outside the class, particularly those designed for English vocabulary development.

Experimental group (EG) received AI-enhanced instruction, while the control group (CG) followed traditional vocabulary instruction. Prior to the study, participants had no experience with AI-based language instruction. Ethical protocols were strictly followed, including obtaining informed consent, ensuring voluntary participation with the right to withdraw, and maintaining confidentiality throughout the research process.

3.3. Instruments and Materials

3.3.1 Oxford Placement Test

In order to choose almost homogeneous participants in terms of English language proficiency, The Oxford Placement Test (OPT) was used as a criterion-referenced test. The OPT test is a widely used assessment instrument comprising 60 components (taking almost 60 minutes) aimed at evaluating overall English proficiency. It provides quick results with Common European Framework of Reference (CEFR) as the indicator of English proficiency (A1-C2). For educators and researchers, OPT provides a reliable and efficient means of placing students at the onset of a course (Allan, 2004).

3.3.2. ESP Course Book

In both groups, the course material utilized in the current research was “Check Your English Vocabulary for Leisure, Travel, and Tourism” authored by Wyatt (2007). Consisting of sixty-two units, this book is a specialized course book designed to help learners improve their vocabulary in hospitality, tourism, and leisure contexts. It features exercises, word games, and practical activities tailored to tourism-specific terms such as hotel bookings, travel arrangement, and customer service. With a clear, structured approach, it covers key terminology, common phrases, and situational dialogs relevant to tourism sector. Altogether, eight units of this book (food terminology, accommodation types, air travel vocabulary, financial terms, transportation vocabulary, money issues, on

the road, and car hire) were covered during the intervention.

3.3.3. Diffit; An AI-powered Platform

Making language learning more inclusive and efficient, Diffit (an AI-powered platform) supports differentiated learning materials tailored to various proficiency levels by simply inputting a topic and the desired complexity. This functionality provided the opportunity to offer a more relevant and engaging learning experience, thereby enhancing students' comprehension of specialized terminology and concepts that were essential in tourism. The possibility to incorporate visuals and graphics responded to the learners' diverse learning styles thus promoting an inclusive learning environment. Adding to the authenticity of the lesson, the participants could type the given topics in the search field. Self-paced learning, language support, test preparation and remediation, and independent exploration are just some of the key features of Diffit.

3.3.4. ESP Vocabulary Test

As the next instrument of the study, tests of vocabulary were administered to evaluate both groups' ability in ESP vocabulary. This assessment focused on the vocabulary content presented in the book "Check Your English Vocabulary for Leisure, Travel, and Tourism" authored by Wyatt (2007). Three vocabulary tests with an approximate time of 20 minutes each were administered to the participants of both groups as pre-test (at the outset of the semester), post-test (in week six), and delayed post-test (in week 12). Each test consisted of 25 fill-in-the-blanks or matching items with each item having one score (See Appendix). The participants' scores at each test could range from 0-25. The tests provided in the course book assessed ESP students' progress of individual units and were designed to align with the topics covered in class and the language taught for this specific course. Pre-test, post-test, and delayed post-test followed the same format and level of difficulty. The pre-test, post-test, and delayed post-test Cronbach's alpha reliability coefficients were 0.78, 0.82, and 0.81, respectively.

In order to ensure the interrater reliability for the ESP vocabulary tests, a university professor with 14 years of experience in teaching English as a foreign language was invited as an independent rater to score the tests as well. As the type of test items required objective scoring (fill-in-the-blanks and matching items), the interrater reliabilities

calculated for the pre-test, post-test, and delayed post-test were 0.91, 0.94, and 0.92 indicating high reliabilities.

3.4. Data Collection and Analysis Procedures

To ensure the homogeneity of the participants, an Oxford Placement Test was administered at the outset of the study. After the exclusion of elementary and intermediate students, 48 participants were randomly assigned to experimental (EG) and control (CG) groups. A Twelve-week intervention was conducted during 2024-2025 academic year in ESP Tourism classes at a university in a southeastern province in Iran.

The researcher (instructor) ensured consistency by implementing identical instructional materials for both groups; however, the experimental group utilized Diffit, an AI-powered platform as the supplementary material during the final 20 minutes of instruction session of each vocabulary unit. This tool provided customized texts with adjustable complexity levels, key vocabulary lists with definitions and contextual examples, multiple-choice and short-answer questions, and multimedia visual aids. Students in EG were also engaged with AI-generated materials (Word, PDF, or Power Point formats) through selected activities so that Diffit generated a text accompanied by a list of key words with definitions and example sentences, multiple-choice questions, and short-answer items. The participants were then directed to select a series of activities from the pre-defined ones and export their generated materials in different formats.

Control group (CG), however, only used the textbook “Check Your English Vocabulary for Leisure, Travel, and Tourism” (Wyatt, 2007). CG students were asked to complete conventional vocabulary exercises on food terminology, accommodation types, air travel vocabulary, financial terms, transportation vocabulary, money issues, on the road, and car hire. Different exercises in the course book focused on the specialized vocabulary that Tourism learners need to understand and use in their profession. It should be mentioned that both groups covered tourism-related vocabulary topics (food and cooking, hotels and accommodation, transport, air travel, money, roads, car, and traffic).

Students of both EG and CG were assessed at week six for their immediate learning gains to check their vocabulary recall as well as week twelve of the semester for their long-term retention effects. Regarding vocabulary recall and retention, the

participants of both groups were assessed through a pre-test, a post-test, and a delayed post-test. In order to address the research questions, a univariate repeated measures ANCOVA was employed to investigate the differences between groups while controlling for pre-test scores. This design allowed for examining both immediate learning gains and long-term retention effects. The following section delves into the results, tabulates the data, and visually illustrates the data analysis pertinent to the study.

4. Results

The research questions were concerned with whether AI-powered instruction had any significant effect on Iranian ESP learners' vocabulary recall and retention. As observed, the mean and standard deviation of word recall and retention variables for the control and experimental groups at the pre-test, post-test, and delayed post-test stages are reported in Table 1.

Table 1.

Descriptive Statistics of Participants' Scores on Pre-test, Post-test, and Delayed Post-test

Group	**Descriptive Index**	**Stages**		
		Pre-test	Post-test	Delayed Post-test
Control	*Mean ± SD*	17.21 ± 3.41	17.54 ± 3.05	17.46 ± 2.84
	Skewness-Kurtosis	-0.15 - 0.73	-0.41 - 0.58	-0.66 - 0.15
Experimental	*Mean ± SD*	17.38 ± 3.59	20.08 ± 2.98	21.04 ± 2.84
	Skewness-Kurtosis	-0.15 - 0.75	-0.02 - 0.74	-0.37 - 0.77

According to Table 1, the mean scores of word recall and retention for both the experimental and control groups changed in the post-test and delayed post-test stages compared to the pre-test stage. These changes indicate that the post-test and delayed post-test scores of students in the experimental group increased in both short-term and long-term word retention.

The research hypotheses were addressed using a univariate repeated measures analysis of covariance (ANCOVA). The use of this analysis requires adherence to certain assumptions, which were examined before conducting the test. The most important

assumptions include: normality of data distribution, homogeneity of variances, homogeneity of regression slopes, and the absence of outliers. To assess the assumption of normality, skewness and kurtosis indices were used. The results in Table 1 show that the skewness and kurtosis indices for word retention fell within the range of -2 to +2, indicating normal distribution. Additionally, the results of the Kolmogorov-Smirnov test (P > 0.05) confirmed the normality of data distribution at the pre-test, post-test, and delayed post-test stages. To examine the assumption of homogeneity of variances, Levene's test was used. The non-significant F value (P > 0.05) indicated that the assumption of homogeneity of variances was also met (Table 2).

To assess the assumption of homogeneity of regression slopes, the significance of the interaction between the pre-test and the grouping variable at the post-test and delayed post-test stages was examined. Given the non-significant F statistic, the assumption of homogeneity of regression slopes was satisfied (F(2,45) = 0.389, P = 0.537). Since all assumptions for the univariate repeated measures ANCOVA were met, the use of this test was justified. Furthermore, Mauchly's test of sphericity was used to examine the assumption of homogeneity of variance-covariance matrices. Given the significance value of Mauchly's test (P = 0.343), the assumption of homogeneity of variance-covariance matrices was also met, and no violation of the statistical model was observed. Additionally, the Greenhouse-Geisser epsilon value indicated that the variance-covariance matrix of the model deviated only slightly from the F statistical model as illustrated in Table 2.

Table 2.

Mauchly's Test of Sphericity for Model Validity

Greenhouse-Geisser Epsilon	Significance Value	Degrees of Freedom	Mauchly's W	Levene's Test P-value	Kolmogorov-Smirnov Test P-value	Stages
0.956	0.343*	2	0.954	0.975* 0.898* 0.868*	0.19* 0.22* 0.29*	Pre-test Post-test delayed Post-test

The results of the univariate repeated measures ANCOVA for the experimental

and control groups at different stages of the study are presented in Table 3.

Table 3.

Univariate Repeated Measures ANCOVA For EG and CG at Different Stages

Source	Sum of Squares	Degrees of Freedom	Mean Square	F	Significance	Effect Size (η^2)
Time	101.431	2	50.715	74.093	0.001*	0.617
Group	158.340	1	158.340	5.623	0.022*	0.109
Time × Group	73.597	2	36.799	53.761	0.001*	0.539

The results in Table 3 show that the effect of the group on the variable word recall among students is significant ($F(1,46) = 5.623$, $P= 0.022$, $\eta^2 = 0.109$). The results indicate that there is a significant difference between the experimental and control groups, with 11% of the variance in the population attributable to the interaction between the dependent variables. Thus, the intervention had a significant effect on improving word recall among students. Additionally, the effect of time on the variable word retention is also significant ($F(2,92) = 74.093$, $P= 0.001$, $\eta^2 = 0.617$). In other words, there is a significant difference in word retention among students at the three stages: pre-test, post-test (short-term memory), and delayed post-test (long-term memory). Furthermore, the interaction effect of time and group on the variable word retention is significant ($F(2,92) = 53.761$, $P = 0.001$, $\eta^2 = 0.539$). This indicates a significant difference in word retention among students in the experimental and control groups across the three stages. The results of the Fisher's Least Significant Difference (LSD) post-hoc test to examine the stability of AI-based vocabulary training on word retention are presented in Table 4.

Table 4.

Results of Fisher's LSD Post-hoc Test for the Stability of AI-based Vocabulary Training on Word Retention

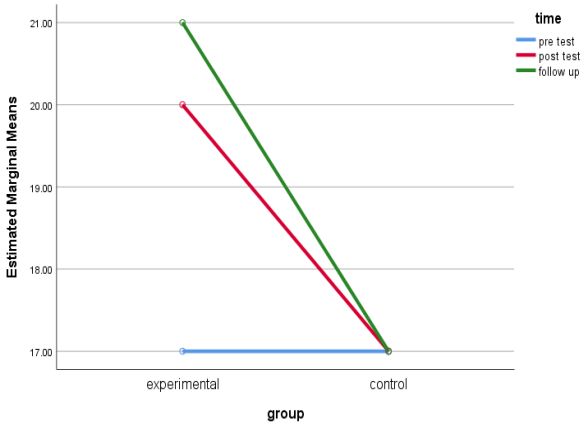
Time	Pre-test	Post-test	Delayed Post-test
Pre-test	-	1.521*	1.958*
Post-test (Short-term Memory)	1.521*	-	0.438*
Delayed			

Post-test (Long-term Memory)	1.958*	0.438*	-

The results of Fisher's LSD post-hoc test in Table 4 show a significant difference between the mean scores of word retention at the pre-test stage and the post-test (short-term memory) stage ($P < 0.05$). Similarly, there is a significant difference between the mean scores at the pre-test stage and the delayed post-test (long-term memory) stage ($P < 0.05$). Moreover, the mean scores of word retention at the delayed post-test (long-term memory) stage were significantly higher than those at the post-test (short-term memory) stage as visually illustrated in Fig 1.

Figure 1.

Comparison of Means of EG and CG Across Three Stages



Therefore, the effectiveness of AI-based vocabulary training on both short-term and long-term memory retention was confirmed and the null hypotheses were rejected.

5. Discussion

The current study demonstrated that AI-enhanced instruction significantly improved both vocabulary recall and retention among ESP learners, corroborating previous research on the efficacy of AI in language learning (Chen & Zhang, 2019; Huang, 2022; Li et al., 2024; Relmasira et al., 2023). These findings align with the established

premise that AI-powered platforms facilitate personalized learning experiences by reducing unnecessary cognitive strain by automating review timing, freeing working memory for deeper processing which entails the enhancement of memory encoding (Sweller, 2019; Sweller et al., 2019).

The superior short-term recall performance in the AI-enhanced group can be attributed to the adaptive and interactive features inherent in AI-based learning systems. Consistent with Sun and Wang's (2020) findings, the integration of multimedia elements in AI platforms appears to heighten learner motivation and attentional focus, both of which are critical for initial vocabulary encoding. Furthermore, the ability of AI systems to contextualize vocabulary within profession-specific scenarios (Wang & Vásquez, 2021) likely enhances depth of processing, leading to more robust short-term retention. This underscores the importance of situated learning in ESP contexts, where lexical items must be mapped onto real-world occupational frameworks to ensure meaningful acquisition.

The long-term memory retention observed in this study support the contention that AI-driven instruction fosters durable lexical retention as a finding that resonates with Zhang and Zou's (2020) work on spaced repetition algorithms. By systematically reintroducing target vocabulary at empirically optimized intervals, AI tools appear to counteract the natural decay of memory traces, facilitating consolidation in long-term storage (Kohnke & Mooorhouse, 2021). Additionally, the dynamic recalibration of content difficulty based on learner performance ensures sustained cognitive engagement without inducing overload, aligning with principles of cognitive load theory (Sweller, 2019; Sweller et al., 2019). This dual mechanism of spaced repetition coupled with adaptive difficulty may explain the experimental group's sustained advantage in the delayed post-test.

The differential success of AI-enhanced instruction may also stem from its capacity to stimulate authentic language use through natural language processing (NLP) capabilities. The AI tool employed in this study (Diffit) operationalized this principle by generating domain-specific texts, vocabulary lists, student activities, and a variety of multimedia supplements thereby narrowing the gap between decontextualized classroom learning and real-world language demands. This contextualization likely promoted deeper

semantic encoding which is instrumental for both recall and retention.

The findings the current study, nevertheless, contradict several recent studies that questioned the efficacy of AI for vocabulary learning. For instance, Bastani et al. (2024) reported that learners using AI platforms performed worse than those using traditional methods possibly due to over-reliance on AI as a learning tool. Likewise, Machin-Mastromatteo (2023) concluded that while AI can support language learning, it often fails to promote deeper retention or meaningful gains in vocabulary among certain learner groups. Trabelsi (2025) also highlighted negative or negligible effects of AI vocabulary retention, attributing to a lack of learner engagement and critical thinking when relying on AI-generated content.

Several factors may explain these contrasting outcomes. First, the current study focused on ESP learners, who may possess higher intrinsic motivation and more targeted vocabulary goals than general language learners, making them more likely to benefit from focused AI activities. Additionally, AI interventions in this study were structured and scaffolded, minimizing the risk of passive learning or over-dependence. Regular feedback and monitoring could have further encouraged active engagement, resulting in stronger vocabulary gains. Lastly, the specific design of AI tasks centered on authentic, discipline-specific contexts may have boosted relevance and retention compared to broader AI interventions described in prior work.

Overall, these findings suggest that when thoughtfully integrated and tailored to learner needs, AI can substantially support ESP vocabulary development even where previous studies found little benefit. Further research should continue exploring which conditions and learner profiles maximize the advantages of AI language learning.

6. Conclusion

This study offers robust empirical evidence supporting the efficacy of AI-enhanced instruction in facilitating vocabulary recall and retention among ESP learners. The findings highlight the transformative potential of AI-driven tools in delivering personalized, adaptive, and contextually rich learning experiences tailored to the special needs of ESP

students. Theoretically, these findings reinforce the applicability of cognitive load theory and depth of processing models to AI-mediated vocabulary learning.

Practically, they emphasize the potential of AI tools to address persistent challenges in ESP instruction in Iran, particularly in settings where traditional methods fail to deliver contextualized, retention-focused training. Most importantly, policymakers should prioritize funding for AI-driven platforms that support adaptive vocabulary instruction, particularly in ESP contexts where domain-specific lexical mastery is critical. This includes providing institutional access to AI tools with spaced repetition, contextual learning, and personalized feedback features. For curriculum designers, the study underscores the value of integration of adaptive AI systems into ESP curricula to support both immediate lexical access and long-term retention. ESP programs should be redesigned to include AI-mediated vocabulary modules, ensuring alignment with ESP communication needs. AI can supplement traditional methods by offering real-world simulations, dynamic assessments, and self-paced learning pathways.

In the current study, several limitations need to be acknowledged. First, the relatively small size (N=48) may constrain the generalizability of the findings, suggesting the need for replication studies within larger and more diverse cohorts. Second, while the study focused specifically on lexical acquisition, it did not examine other critical dimensions of language proficiency such as grammatical accuracy or oral fluency. Future investigations could productively explore the impact of AI on these under examined competencies. Additionally, the retention period assessed in this study was limited to twelve weeks; longitudinal research spanning extended durations would help ascertain the durability of AI-enhanced learning outcomes.

To advance this line of inquiry, several promising research directions emerge. First, the relationship between AI-mediated instruction and learner autonomy warrants systematic investigations. Although AI platforms provide individualized learning pathways, the degree to which they foster or constrain self-regulated learning strategies remains an open question. Second, comparative studies examining blended learning models where AI tools are strategically integrated with conventional pedagogical approaches could yield valuable insights for optimizing instructional design. Finally,

qualitative explorations of learner experiences with AI systems may illustrate the affective and cognitive processes underlying vocabulary acquisition in technology-enhanced environments. Future research should explore the practicality of such interventions across diverse ESP domains and learner populations.

In conclusion, this study makes a substantive contribution to the burgeoning literature on AI in second language acquisition, particularly within ESP contexts. By empirically validating the benefits of AI-enhanced vocabulary instruction, it underscores the affordances of AI to address persistent challenges in specialized language education. However, as with any emerging pedagogical innovation, these findings should be interpreted as preliminary rather than definitive. Continued interdisciplinary research spanning applied linguistics, educational technology, and cognitive science will be essential to fully realize transformative potential of AI while addressing its current limitations. The present study thus serves as both a foundation for future inquiry and a call for more nuanced investigations at the intersection of artificial intelligence and language pedagogy. It not only advances the discourse on AI in ESP education but also invites broader reflection on how emerging technologies can reshape Teaching English as a Foreign Language (TEFL) by balancing efficiency with pedagogical depth—ensuring that AI serves as a tool for meaningful language learning rather than a mere technological quick fix.

References

Alaqlobi, O., Alduais, A., Qasem, F., & Alasmari, M. (2024). Artificial intelligence in applied linguistics: A content analysis and future prospects. *Cogent Arts and Humanities, 11*(1), 1-19.

Alhusaiyan, E. (2025). A systematic review of current trends in artificial intelligence in foreign language learning. *Saudi Journal of Language Studies, 5*(1), 1-16. https://doi.org/10.1108/SJLS-07-2024-0039

Allan, D. (2004). *Oxford placement test 2: Test pack.* Oxford University Press.

Al Zahrani, S. M., & Chaudhary, A. (2022). Vocabulary learning strategies in ESP context: Knowledge and implication. *Arab World English Journal, 13*(1), 382-393. https://doi.org/10.24093/awej/vo113no1.25

Bastani, M., Zhang, A, & Lam, C. (2024). Effects of AI tutors on high school English learners' vocabulary outcomes: An experimental study. *Journal of Educational Technology, 39*(2), 112-128.

Chen, X., Zou, D., Xie, H., & Cheng, G. (2020). Twenty years of personalized language learning: Topic modeling and knowledge mapping. *Educational Technology & Society, 23*(1), 59-72. https://doi.org/10.30191/ETS.202001_23(1).0005

Chen, Y., & Zhang, J. (2019). AI-driven adaptive learning systems for ESP vocabulary acquisition. *Computer Assisted Language Learning, 32*(5-6), 567-589.

Coxhead, A. (2022). *Vocabulary and English for specific purposes research: Quantitative and Qualitative Perspectives.* Routledge.

Cui, Y. (2024). Application of artificial intelligence technology in college English vocabulary teaching. *International Journal of New Developments in Education, 6*(7), https://doi.org.10.25236/IJNDE.2024.060732

Dudley-Evans, T., & St John, M. J. (1998). *Developments in English for specific purposes.* Cambridge University Press.

Gholami, J., & Khosravi, H. (2022). Innovative approaches in teaching English for specific purposes: A review of current practices. *International Journal of Educational Research, 112*, 101-115.

Hao, Y., & Liu, M. (2022). Cognitive load and academic performance in the age of AI: The role of AI tools in academic writing. *Computers & Education*, 185, 104222.

Hayati, A. M., & Jalilifar, A. (2009). The impact of ESP on Iranian EFL learners' vocabulary learning. *Journal of English for Academic Purposes, 8*(1), 3-12. https://doi.org/10.1016/j.jeap.2008.12.002

Huang, F., & Zou, B. (2021). English speaking with artificial intelligence (AI): The roles of enjoyment, willingness to communicate with AI, and innovativeness. *Computers in Human Behavior, 159*, https://doi.org/10.1016/j.chb.2024.108335

Huang, X., Zou, D., Cheng, G., & Xie, H. (2021). A systematic review of AR and AI in language learning. *Computer Assisted Language Learning, 36*(2), 1-27. https://doi.org/10.1080/09588221.2021.1930054

Huang, M. (2022). Artificial intelligence in education: Transforming language learning through NLP. *Educational Technology & Society*, 25 (3), 12-24.

Hwang, G.-J., Xie, H., Wah, B. W., & Gašević, D. (2020). Vision, challenges, roles and research issues of artificial intelligence in education. *Computers and Education: Artificial Intelligence, 1*, Article 100001. https://doi.org/10.1016/j.caeai.2020.100001

Kohnke, L., & Moorhouse, B. L. (2021). Vocabulary learning with AI-powered flashcards: A case study in ESP. *Language Learning & Technology, 25*(1), 56-73.

Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for language teaching and learning. *RELC Journal, 54*(2), 1-12. https://doi.org/10.1177/00336882231162868

Kundu, A., & Parida, L. (2022). Digital literacy and student engagement with AI in ESL classrooms. *International Journal of Educational Research Open*, *3*, 100-126

Li, B., Bonk, C., Wang, C., & Kou, X. (2024). Reconceptualizing self-directed learning in the era of generative AI: An exploratory analysis of language learning. *IEEE Transactions on Learning* Technologies*,* 17. https://doi.org/10.1109/TLT.2024.3386098

Lin, M., & Vuono, A. (2019). AI-driven vocabulary learning in ESP: A longitudinal study. *Journal of Computer-Assisted Learning, 35*(4), 456-470.

Lin, C. C., & Huang, A. Y. (2021). AI-assisted vocabulary learning in ESP: A quasi-experimental study. *Interactive Learning Environments, 29*(4), 567-582.

Liu, T. (2017). Developing an English mobile learning attitude scale for adult learners. *Journal of* Educational *Technology, 45* (3), 424-435. https://doi.10.1177/0047239516658448

Lu, X., & Li, G. (2020). The impact of AI-powered vocabulary apps on L2 learners' vocabulary retention. *Journal of Educational Technology & Society, 23*(3), 78-92.

Lyu, B., Lai, C., & Guo, J. (2024). Effectiveness of chatbots in improving language learning: A meta-analysis of comparative studies. International Journal of Applied Linguistics, 35, 834-851. https://doi.org/10.1111/ijal.12668

Machin-Mastromatteo, J. D. (2023). Artificial intelligence in English language teaching: The good, the bad, and the ugly. *RELC Journal, 54*(1), 191-208.https://doi.org.10.1177/00336882211053459

Macia. A. (2012). The role of technology in teaching languages for specific purposes courses. *The Modern Language Journal, 96,* 89-104.

Makhmudova, F. S., & Mashrapova, S. X. (2024). Importance of vocabulary in English for specific purposes teaching and learning. *TADQIQOTLAR.UZ*, *40* (4), 89-95.

Mayer, R. E. (2021). *Multimedia learning (3rd ed.)*. Cambridge University Press.

Medina, M., Mejia, R. A., Mejes, D. M., & Bustamante, M. (2024). Self-reported preparedness and factors influencing AI tools integration in teaching among Masters' degree students in a selected teachers' education college. *International Journal for Multidisciplinary Research, 6* (1), 1-34.

Mostafaei Alaei, M., & Ershadi, A. (2016). ESP program in Iran: A stakeholder-based evaluation of the program's goal, methodology, and textbook. *Issues in Language Teaching, 5*(2), 279-306.

Nation, I. S. P. (2013). *Learning vocabulary in another language*. Cambridge University Press.

Pavlik, P. I., & Anderson, J. R. (2008). Using a model to compute the optimal schedule of practice. *Journal*

of Experimental Psychology: Applied, 14(2), 101–117. https://doi.org/10.1037/1076-898X.14.2.101

Pham, T. T., Nguyen, L. A. D., Dang, H. M., & Le, T. T. P. (2024). Exploring tertiary Vietnamese EFL students' engagement in vocabulary learning through the use of an AI tool. *Proceedings of the Asia CALL International Conference, 4,* 129-149. https:// doi.org10.54855.paic.2340

Pokupec, M., & Njerš, D. (2014). Constructive alignment in assessing vocabulary in business English study course. *The 7th international language conference on the importance of learning professional foreign language for communication between cultures,* Faculty of Economics and Business of the University of Maribor.

Relmasira, S. C., Lai, Y. C., & Donaldson, J. P. (2023). Fostering AI literacy in elementary science, technology, engineering, art, and mathematics (STEAM) education in the age of generative AI. *Sustainability, 15*(18), Article 13595. https://doi.org/10.3390/su151813595

Rezaei Ali Kamar, F., Mojallal, M., Ghahremani, J. (2021). Identifying high school technological competency-based curriculum: A qualitative study. *Curriculum Research, 2* (1), 23-34.

Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. Palgrave, Macmillan.

Selvi, V. T., & Thirumoorthi, K. (2024). Applications and analyzes of artificial intelligence in enhancing English vocabulary. *Recent Research Reviews Journal, 3*(2), 334-345. https://doi.org.10.36548/rrrj.2024.2.002

Silitonga, L. M., Wiyaka, & Prastikawati, E. F. (2024). Boosting students' ESP vocabulary by utilizing AI chatbot. *Eternal: English Teaching Journal, 15* (2), 275-283. https://doi.org/10.26877/eternal.V15i2.605

Stockwell, G., (2013). Technology and motivation in English-language teaching and learning. In E. Ushioda (Ed.), *International perspectives on motivation: Language learning and professional challenges* (pp. 156-175). London: Palgrave Macmillan.

Sun, Y., & Wang, L. (2020). AI-based personalized vocabulary learning in ESP: A quasi-experimental study. *Journal of Educational Computing Research, 58*(4), 789-812.

Sussmann, M. (2024). AI tools for the four skills in English language learning. *TESOL Connections*, 2-4.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12*(2), 257-285. https://doi.org/10.1207.s15516709cog1202-4

Sweller, J. (2019). Cognitive load theory and educational technology. *Educational Technology Research and Development*, 67 (1), 1-16.

Sweller, J., van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review, 31*(2), 261–292. https://doi.org/10.1007/s10648-019-09465-5

Trabelsi, W. (2025). The good, the bad, and the ugly science of AI in education: A critical review. *Computers & Educational Review, 13*(1), 55-69.

Tuan, V. (2022). Collaborative writing and the role of AI: A classroom-based study. *Asian EFL Journal*, 24(16), 112-131.

Valencia, H. G., Gomez, G. D. I., Idarraga, M. A., & Villaquira, M. R. (2020). Integrating the virtual platform memrise as a teaching tool for English vocabulary in foreign language students of second semester at a university. *Revista Latinoamericana de Estudios Educativos, 16* (1), 259-284.

Wang, S., & Vásquez, C. (2021). The role of chatbots in ESP vocabulary learning: A quasi-experimental study. *ReCALL, 33*(2), 145-162.

Woodrow, L. (2017). Vocabulary and English for specific purposes. In *Introducing course design in English for specific purposes* (pp. 69-85). Routledge.

Wyatt, R., (2007) .*Check your English vocabulary for leisure, travel, and tourism* (second edition). A & C Black Publishers.

Zhang, R., & Zou, D. (2020). AI-based gamification in ESP vocabulary learning: A mixed-methods study. *System*, *91*, 102-117.

Zou, D., & Xie, H. (2019). AI-powered vocabulary learning in ESP: A case study of tourism students. *Journal of Language Teaching and Research, 10*(5), 1023-1032.

Appendix

Sample Items from Vocabulary Test

From the book "Check Your English Vocabulary for Leisure, Travel, and Tourism" authored by Wyatt (2007)

-Application form	- driving license	- hotel voucher	- food hygiene certificate
-Flight coupon	- claim form	- ID card	- customer satisfaction questionnaire
-Baggage check	- exit visa	- landing card	- health declaration form
-Boarding pass	- form E 111	- passport	- certificate of seaworthiness
- transit visa	- ticket	- multiple-entry visa	- property irregularity report
- revalidation sticker	- work permit	- rental agreement	- certificate of airworthiness
- travel insurance	-vaccination certificate	- docket	- clearance certificate
- landing card	- receipt	- travel voucher	

1. Your flight to Tokyo has a 12-hour layover in Moscow. If you want to leave the airport and visit the city, you will need a ---------, which you can get from the Russian embassy before you leave.
2. Ladies and gentlemen, We will shortly be arriving in Athens. Non-EU citizens will need to fill in a ---------- before going through immigration, and we will be handing these out now.
3. This is an advance purchase, promotional, round trip, off-peak, non-endorsable, non-transferable, non-refundable, economy class, maximum stay, open-ended ----------. Do you think you can remember that?
4. At the airport, go to the check-in, show them your ticket, give them your baggage and collect your ----------, which will show your seat number, boarding time and gate number.
5. At the reception, give the receptionist your ----------. This shows that you have booked and paid for your room. It also shows that breakfast is included in the price.
6. When a customer buys a package holiday, the tour operator will often send ---------s to the airline, the hotel, etc. to pay for the holiday.
7. European Union residents visiting other European Union countries can get free or reduced-cost medical assistance if they have a ---------- with them.
8. You should always have ---------- when you go on a trip, just in case you lose something valuable, have something stolen or need medical treatment.
9. Some countries will not let foreigners in if their ---------- is valid for less than six months. If this applies to you, you will need to fill in an ---------- for a new one.
10. There are two parts to your airline ticket: the ----------, which the check-in staff keep, and the receipt, which you keep with you.
11. When you hire a car, it is very important to read the ---------- very carefully before you sign it. You will also need to show your ----------.
12. In a lot of countries, you need to carry an ---------- at all times, so that you can prove who you say you are.
13. Before you start a job in another country, it is usually essential to obtain a ----------.

14. All aircrafts must have a ---------- before they are allowed to fly. Similarly, a ship must have a ---------- before it is allowed to sail.
15. Goods that go from one country to another have to be accompanied by a ---------- to show that they have been passed by customs.
16. Some countries may require foreign visitors to have a ---------- that shows they are immune to certain diseases that they could catch in that country before they will let them in. others may ask to see a ---------- to show that visitors are in good health and free from contagious diseases.
17. If an airline loses a passenger's baggage, they will ask him to fill in a ----------, describing the item of baggage and its contents. The passenger should give this form, together with his ---------- (which shows that his baggage was checked in by the airline) to a member of the ground crew.
18. In many countries, a restaurant needs to have a ---------- to show that it meets national standards of cleanliness.
19. Travel companies often ask their guests to fill in a ---------- at the end of their holiday so that they can find out if they need to make any changes or improvements to the way they operate.
20. If you have something stolen while on holiday and want your insurance company to replace it, you will need to fill in a ---------- describing what was stolen and how much it was worth.
21. When you buy something, you should always ask for, and keep the ---------- in case you need to return it.
22. When the hotel takes a delivery of something, it is important to check the accompanying ---------- to make sure that everything the hotel ordered is there.
23. If an airline passenger decides to change her flight times or another aspect of her flight, it is not always necessary to give her a new ticket. Sometimes a ---------- is placed on her original ticket to show that a change has been made.
24. This is a ----------, which means that you can enter and leave the country as many times as you like during a specific period.
25. Some countries require travelers to have an ---------- before they let them leave the country.