

Evaluating the role of artificial intelligence (AI) in enhancing vocabulary retention in foreign language acquisition



Nisar Ahmad Koka¹, Basim Kana'an¹, Sheeba Hassan², Javed Ahmad¹, Nusrat Jan², Mohamad A. Khasawneh³,
 Mohammad A. Tashtoush^{3,4,*}

¹Department of English, College of Languages and Translation, King Khalid University, Abha, Saudi Arabia

²Department of Linguistics, University of Kashmir, Jammu and Kashmir, India

³Faculty of Education and Arts, Sohar University, Sohar, Oman

⁴Department of Basic Sciences, Al-Huson University College, Al-Balqa Applied University, Salt, Jordan

ARTICLE INFO

Article history:

Received 13 October 2025

Received in revised form

2 March 2026

Accepted 9 March 2026

Keywords:

Artificial intelligence

Vocabulary retention

Language learning

Personalized learning

Student engagement

ABSTRACT

This study examines the impact of Artificial Intelligence (AI) tools on vocabulary retention among foreign language learners in Saudi Arabia. A quasi-experimental design was used with 150 students divided into two groups: an experimental group that used AI-based learning tools and a control group that followed traditional teaching methods. The results show that students who used AI tools achieved significantly higher vocabulary retention and showed greater engagement than those in the control group. Moreover, students in the experimental group reported spending more time on learning activities and expressed higher satisfaction with their learning experience. These findings suggest that AI tools can improve traditional language education by providing personalized learning support for vocabulary acquisition. The study also highlights the potential long-term benefits of integrating AI into language learning and recommends further research to better understand its broader role in educational development.

© 2026 The Authors. Published by IASE. This is an open access article under the CC BY-NC-ND license (<https://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Artificial intelligence (AI) enters education through language learning programs to generate innovative teaching approaches that improve student educational quality. Learning new foreign language (FL) vocabulary presents a significant retention challenge because students frequently have difficulty maintaining received words in their long-term memory storage. Traditional educational approaches combining rote memorization and textbook exercises control classroom settings today, but Fu (2009) said these methods fail to deliver personalized learning experiences for diverse students. The assessment of AI-based solutions must occur to understand their potential for solving educational problems and developing better learning results.

The combination of AI technology with intelligent tutoring systems and gamified learning applications

enables personalized, real-time student feedback that allows learners to select individualized learning routes. Advanced algorithms integrated into these tools implement adaptive learning methods along with spaced repetition to enhance vocabulary retention capabilities. Lindsey et al. (2014) proved that adaptive forgetting curves enhance student vocabulary review efficiency, thus leading to better vocabulary retention across extended periods. AI-based educational methods create challenges to their use across diverse cultural zones and linguistic backgrounds. The tools show promise, but their design foundations stem from Western educational theory, which differs from the learning needs and learning styles of Saudi Arabian students (O'Brien and Toms, 2008; Halabi, 2017).

The learning environment in Saudi EFL positions itself perfectly to understand the influence of AI tools for vocabulary attainment. Limited real-world exposure to English in Saudi Arabia causes second-language learners to struggle with their vocabulary learning despite widespread classroom instruction of English as a second language. The Saudi educational system heavily depends on teacher-centered teaching supplemented by standardized curricula, according to Helsing et al. (2008) and Farmer et al. (2021), whose framework does not elicit sufficient adaptive modification for unique

* Corresponding Author.

Email Address: tashtoushzz@su.edu.om (M. A. Tashtoush)

<https://doi.org/10.21833/ijaas.2026.03.010>

Corresponding author's ORCID profile:

<https://orcid.org/0000-0002-2436-8155>

2313-626X/© 2026 The Authors. Published by IASE.

This is an open access article under the CC BY-NC-ND license

(<https://creativecommons.org/licenses/by-nc-nd/4.0/>)

student needs. AI-based intervention systems possess adaptive properties that enable second language learners to participate in meaningful interactive language use activities, making up for this knowledge gap. These interventions need effective cultural and contextual nuance integration in their design to show optimal results, yet existing studies have failed to focus on this vital aspect.

Laboratory research about AI tools demonstrates improved vocabulary acquisition together with better retention, yet evidence shows variation across different studies. Aldowsari and Aljebreen (2024) showed that AI-based language apps enhance Chinese EFL learners' vocabulary retention; however, such findings require more validation in Saudi Arabia. Using technology in learning leads to ethical as well as logistical challenges, particularly in matters related to data privacy, together with the availability of AI tools among disadvantaged students. Studies by Albalwi (2023) and Aloufi et al. (2024) indicated that educational AI tools enhance instructional results, yet their successful deployment needs solid technical equipment and competent trainer support, which Saudi educational organizations presently lack equitably.

The discussion requires a transition to teaching approaches that allow users to achieve optimal advantages from AI systems. Al-Faqih (2015) establishes that Saudi Arabia's conventional EFL classrooms base their instruction on rote learning techniques and grammatical precision, yet these approaches disagree with the exploration-driven aspects of AI-based educational spaces. Some educators show opposition to adopting innovative teaching technologies because these platforms remain outside their professional understanding or doubt their educational impact. A comprehensive solution for these challenges needs an understanding of how technology works together with teaching methods. Educators function as essential mediators between students and AI-driven learning platforms by providing effective guidance for tool usage, along with teaching students critical thinking and independent learning abilities.

Although AI presents promising advantages for vocabulary retention, the critical analysis demonstrates the necessity of maintaining equilibrium in this realm of applications. Research performed by Beirat et al. (2025) demonstrates AI's value in language education, yet warns against repetitive technology usage when students need direct assistance from human instructors. The ability for AI technologies to replicate social and cultural aspects of language learning remains doubtful because they do not produce authentic communicative competence development. Gamified vocabulary applications improve student engagement; however, they often prove insufficient for teaching language use in actual environments, which demand intricate socio-cultural context understanding.

The problem of the study is that foreign language (FL) students in Saudi Arabia face ongoing

difficulties sustaining vocabulary knowledge despite their limited external English exposure. Traditional educational approaches struggle to accommodate different student learning requirements because they depend primarily on memorization and standard practices that are not personally responsive or engaging for learners. Such teaching methods produce poor learning retention, which makes it difficult for students to store vocabulary accurately in long-term memory and deploy it in regular conversations. AI tools have increasingly become popular in language learning by providing individualized yet interactive educational experiences, but experts have investigated their effect on vocabulary retention in Saudi EFL classrooms insufficiently. This study investigated whether AI platforms improve vocabulary memory retention performances of Saudi Arabian FL students based on collected experimental evidence for classroom implementation. Questions of the study are as follows:

- Q1.** What is the impact of AI tools on vocabulary retention rates among foreign language learners in Saudi Arabia?
- Q2.** How do AI-based vocabulary learning interventions compare with traditional teaching methods in terms of improving long-term retention?
- Q3.** What factors influence the effectiveness of AI tools in enhancing vocabulary retention for Saudi EFL learners?

Research findings will support educators, together with researchers and policymakers across Saudi Arabia and international contexts. The research collects actual data about how AI platforms assist in tackling language educators' core retention problem with foreign terms. Through its demonstration of AI-driven learning potential, the research establishes practical guidelines for incorporating AI technology into EFL lesson environments. These findings enrich academic knowledge about how technological integration affects second language instruction, particularly in culturally defined teaching environments. The study reveals important implications for educators to use integrated teaching techniques that adjust learning plans based on different student needs. These findings also offer valuable insights for policymakers, guiding investment toward strengthening educational technology infrastructure and providing targeted professional development to support the effective implementation of AI tools. The research illustrates how AI technology generates revolutionary advancements in educational methods, which lead to customized educational experiences featuring innovative instructional techniques.

A period of eight weeks was established to comprehend how AI tools affect Saudi EFL learners' ability to maintain their vocabulary. Real-world testing lasted eight weeks to let students experience AI interventions fully while researchers could assess sustained knowledge acquisition. Throughout the

study, multiple tests were given before teaching began to measure immediate retention and again after instruction, and later to evaluate retention effects over time. A broad range of learners from various educational settings participated in this study to ensure that results applied universally to multiple populations. The research used the best available AI learning tools, which met accessibility requirements and provided value for vocabulary acquisition through adaptive feedback systems and gamified interfaces with spaced repetition functionality. The researchers designed a study timeline in compliance with ethical conventions, which provided a systematic evaluation of the research goals.

Several constraints affect the understanding we gained through this research regarding AI tools' ability to maintain linguistic memory. The representative sample size lacks complete representation of Saudi EFL learners, which reproduces limitations in finding applicability across different student groups. The investigation focused solely on AI tools without considering the potential diversity of current technological capabilities. AI applications show varying effectiveness because of their different designs and features, and user interface elements, so more research must evaluate these AI solutions against one another. The eight-week duration proved informative for identifying short-term and some long-term retention effects; however, it might fail to demonstrate AI-based intervention persistence across extensive time periods. This research took place within Saudi EFL zones, so its conclusions may not transfer rapidly to other cultural and academic institutions. Technical obstacles, including participant differences in AI tool familiarity, contributed to study outcomes, which demonstrate that teacher training and technical assistance could improve future system implementations.

2. Literature review

The way learners hold onto new vocabulary plays a vital role in their language acquisition process, determining both their communication abilities and their level of mastery. However, a pervasive challenge in FL education is the rapid decay of newly acquired vocabulary from memory. These cognitive principles highlight that without consistent reinforcement, learned information, particularly isolated lexical items, is quickly forgotten over time. Recognizing this inherent challenge, researchers are increasingly dedicating efforts to leverage technology, particularly AI, to deliver customized, interactive, and adaptive learning environments (Sajja et al., 2025). AI-powered solutions, including intelligent tutoring systems, gamified platforms, and adaptive learning applications, are designed to accommodate specific learner requirements and directly address the weaknesses inherent in traditional teaching systems, especially concerning long-term vocabulary retention.

Central to AI tools' success in maximizing vocabulary retention is their sophisticated utilization of Spaced Repetition Systems (SRS). These systems are scientifically grounded in the Ebbinghaus Forgetting Curve, which postulates that optimal memory recall is achieved by reviewing information at progressively longer intervals just before it is likely to be forgotten. AI-driven SRS algorithms precisely calculate these strategic learning intervals for each word and learner, thereby fortifying memory retention more efficiently and mitigating cognitive overload compared to uniform review schedules. As demonstrated by Hatamleh et al. (2025), AI-based tools, by dynamically adapting to each learner's unique forgetting patterns, deliver a noticeable boost in long-term retention for language learners. This adaptive approach ensures that instructional effort is concentrated where it is most needed, enhancing learning efficiency and effectiveness.

Beyond optimized scheduling, AI-based vocabulary learning tools significantly promote learner engagement through their interactive and gamified designs. Such elements are crucial for harnessing intrinsic motivation, aligning with theories like Self-Determination Theory, where autonomy, competence, and relatedness drive sustained engagement. According to Alenezi (2023), learners achieve better retention through gamified AI applications, which not only enhance active engagement and foster feelings of accomplishment but also encourage deeper cognitive processing. This interactive design, which helps students connect vocabulary items to real-world scenarios, directly addresses principles of the Depth of Processing Hypothesis. This hypothesis suggests that the more deeply and elaborately information is processed, e.g., through semantic connections, contextual use, or problem-solving rather than rote memorization, the more likely it is to be retained. AI tools facilitate this by providing varied exercise types, contextual examples, and even AI-driven storytelling or chatbot interactions, which move learners beyond superficial memorization techniques towards more meaningful encoding and retrieval practices, thereby enhancing the quality of memory traces.

While research broadly confirms the effectiveness of AI language learning tools, their optimal implementation and impact are significantly modulated by specific contextual elements and cultural characteristics. The Saudi EFL learning environment, for instance, presents unique challenges, including limited native English exposure and a prevailing tradition of teacher-led instruction that may not fully align with the learner-centric and exploratory nature of AI-based platforms (Halabi, 2017). This highlights the critical need to consider not only the technological capabilities of AI but also the socio-cultural and pedagogical frameworks into which they are integrated, ensuring that instructional approaches leverage AI effectively while being sensitive to local learning needs and styles.

New research through comparative analysis continues to uncover important findings about the effectiveness of AI tools for educational situations. A study by [Chen et al. \(2022\)](#) demonstrated that AI vocabulary learning modules surpassed traditional methods because learners achieved better retention results and enjoyed higher satisfaction levels. According to [Teng and Reynolds \(2024\)](#), personalized feedback coupled with contextualized practice emerged through meta-analyzing corpus-based tool effects on vocabulary learning as essential factors in language acquisition.

Similar investigations have noted several restrictions that researchers believe exist when using AI-based vocabulary learning systems. As [Srinivasa et al. \(2022\)](#) explained, AI tools deliver many benefits, yet their ability to influence learning can be reduced by hardware glitches and students who lack technical competency, as well as teacher resistance toward AI education methods. To achieve equitable access to AI-driven educational resources in educational institutions, you need to resolve both data privacy and digital divide ethical issues.

The academic research demonstrates how AI tools enable large-scale improvements for vocabulary learning, especially for countries such as Saudi Arabia, which need better teaching systems. AI technologies utilize adaptive learning algorithms with gamified elements alongside personalized feedback for enhancing vocabulary acquisition among learners, along with their vocabulary retention abilities. The successful execution of these tools depends on a proper understanding of contextual conditions alongside strong infrastructure capabilities and continual educator involvement. Research on existing work shows updated insights regarding how AI tools affect vocabulary retention in Saudi EFL learners while expanding technology-enhanced FL teaching knowledge.

3. Methods

A quantitative research design analyzed how AI tools impact vocabulary retention for FL learners throughout Saudi Arabia. The framework provided a solid foundation to acquire dependable data that answered the research questions completely. The analysis describes both research participants together with their data collection procedures, along with measurement tools, which enabled statistically sound final data interpretation.

The researchers gathered data from 150 Saudi EFL learners from three public universities throughout Saudi Arabia. Specific participant selection through stratified random sampling produced diverse populations based on academic levels and proficiency levels, along with male and female representation. It was a must for the participating learners to attend intermediate English classes with foundational experience using digital tools to effectively utilize AI-based applications for the study. The research used 80 female and 70 male

participants with a mean age of 20.6 years. The selection roles included learners from different socioeconomic backgrounds as well as urban and rural communities to produce research findings that applied across wider demographic segments.

The research lasted eight weeks throughout an academic term. Participants were divided into two groups: an experimental group (n=75) and a control group (n=75). Participants in the experimental group learned vocabulary through an AI-based application while the control group continued using established textbook activities and subject matter instruction. All participants received instruction on 100 vocabulary words derived from their school curriculum.

The data collection process involved three key phases: pre-test, intervention, and post-test. Researchers gave all participants an initial baseline vocabulary knowledge test before the study began. The experimental group used the AI tool for daily sessions of 20-30 minutes throughout the intervention phase, although the control group received an equivalent period of traditional learning. The researchers performed an immediate post-test to measure short-term retention at the end of the four-week study period. The long-term retention evaluation through a delayed post-test took place four weeks after the initial assessment. The timing of the identical items for post-testing allowed researchers to guarantee consistent retention measurement.

A learning instrument designed for artificial intelligence purposes, with features of spaced repetition, adaptive learning with feedback, and gamification, was chosen as the treatment instrument for the experiment group. This is an innovative tool because it combines evidence-based principles of learning. It is a personalized learning tool because it uses an adaptive learning function with a continuous mechanism monitoring learners' performance concerning accuracy rates and reaction times, adjusting the level of difficulty of new words introduced with varying exercise types (match, fill-the-blank, and constructing sentences) and learning paths automatically in a split second. This tool is a fine example of a learning device with a robust spaced repetition algorithm like the SM-2 algorithm, designed automatically for each learner's forgetting curve, so that words can be repeated according to a progressively longer period of optimal times for better retention rates, with a strong emphasis on Gamification principles for increased motivation. Another significant design quality is the automatic tracking feature of participant activity, including duration, accuracy levels, and navigation of educational content, which is a fundamental factor for measuring learning activity and learning outcomes throughout this experiment.

Analyses of the information gathered took place through SPSS version 28. The researchers computed descriptive statistics from pre-test, post-test, and delayed post-test data using group means and standard deviation measurements. The researchers applied a paired-samples t-test to evaluate pre-test

and post-test score evolutions, relying on internal group data for understanding vocabulary maintenance across durations. A comparison of AI tool effectiveness versus conventional methods occurred through an independent-samples t-test that analyzed experimental group performance against the control group. Long-term retention was measured through repeated-measures ANOVA that examined score variations between pre-test and post-test and delayed post-test for each experimental group. The research team conducted post-hoc tests with Bonferroni adjustments to detect changes in patterns that emerged during the study. A regression analysis used information from usage logs to study how students engaged with the AI system while revealing its effectiveness through retention score assessment.

4. Results

Table 1 shows that the p-value for all tests the p-values for all tests indicate no significant differences in variances between the experimental and control groups. The results validate the assumption of homogeneity of variances for further parametric analyses. Table 2 shows that the p-value confirm that the p-values confirm that the data for both groups at all stages of testing (pre-test, post-test, delayed post-

test) were normally distributed, meeting the assumption required for parametric tests.

The data displayed in Table 3 shows equivalent initial pre-test performance levels among the two groups. The experimental group scored notably higher on post-test assessments in all stages, thanks to the AI tool, which led to improved retention performance. From the data in Table 4, both groups demonstrated significant improvements from the pre-test to the post-test (p = 0.000 for all comparisons). However, the experimental group displayed larger effect sizes, suggesting that the AI tool substantially enhanced vocabulary retention. According to Table 5, the experimental group significantly outperformed the control group in both the immediate and delayed post-tests (p = 0.000 for both). The large effect sizes indicate that the AI tools had a strong positive impact on vocabulary retention.

Table 1: Homogeneity test (Levene’s test for equality of variances)

Test	F	P-value	Interpretation
Pre-test scores	0.832	0.451	Variances are equal across groups
Immediate post-test scores	1.124	0.298	Variances are equal across groups
Delayed post-test scores	0.762	0.486	Variances are equal across groups

Table 2: Normality test (Shapiro-Wilk test)

Group	Test	Statistic	P-value	Interpretation
Experimental group	Pre-test scores	0.978	0.120	Data is normally distributed
	Post-test scores	0.982	0.078	Data is normally distributed
	Delayed post-test	0.975	0.142	Data is normally distributed
Control group	Pre-test scores	0.981	0.095	Data is normally distributed
	Post-test scores	0.973	0.167	Data is normally distributed
	Delayed post-test	0.970	0.181	Data is normally distributed

Table 3: Descriptive statistics

Test	Group	Mean (M)	Standard deviation (SD)	N
Pre-test scores	Experimental group	58.24	9.12	75
	Control group	57.89	9.34	75
Immediate post-test	Experimental group	78.45	8.25	75
	Control group	68.32	9.47	75
Delayed post-test	Experimental group	74.21	8.87	75
	Control group	63.15	9.92	75

Table 4: Paired-samples t-test (within-group comparisons)

Group	Test pair	Mean difference (MD)	T-value	P-value	Effect size (Cohen’s d)
Experimental group	Pre-test vs post-test	20.21	11.54	0.000	1.32
	Pre-test vs delayed post-test	15.97	9.87	0.000	1.15
Control group	Pre-test vs post-test	10.43	7.89	0.000	0.86
	Pre-test vs delayed post-test	5.26	4.12	0.000	0.49

Table 5: Independent-samples t-test (between-group comparisons)

Test	Mean difference (MD)	T-value	P-value	Effect size (Cohen’s d)
Immediate post-test	10.13	5.76	0.000	0.97
Delayed post-test	11.06	6.02	0.000	1.02

The results from repeated-measures ANOVA in Table 6 demonstrated a significant impact on both group (F = 58.76, p = 0.000) and time (F = 79.54, p = 0.000) alongside a significant group-time interaction (F = 26.83, p = 0.000). Large effect sizes become evident through partial eta squared values because the experimental group demonstrated better performance throughout time than the control

group. Engagement evaluations from the experimental participants yielded superior outcomes than their control counterparts (Table 7), and these differences reached statistical significance (p = 0.000). The analysis shows that AI-assisted educational tools create an environment that motivates students to participate actively in their learning activities.

Table 6: Repeated-Measures ANOVA

Source	Sum of squares	df	Mean square	F-value	P-value	η^2 (Partial eta squared)
Group	2456.89	1	2456.89	58.76	0.000	0.38
Time	3921.43	2	1960.71	79.54	0.000	0.51
Group \times time	1321.67	2	660.84	26.83	0.000	0.24
Error	5268.94	296	17.80			

Table 7: Engagement levels in learning

Item	Experimental group (M)	SD	Control group (M)	SD	T-value	P-value
The learning process was engaging	4.65	0.48	3.42	0.87	9.76	0.000
Activities increased my interest in vocabulary	4.78	0.40	3.33	0.91	11.14	0.000
I felt motivated to participate actively	4.56	0.57	3.21	0.79	10.22	0.000

Self-reported questionnaire results; 5-point Likert scale: 1 = Strongly disagree, 5 = Strongly agree

Across every word category, the experimental group maintained more consistent recognition levels. Retention success using the AI-supported method demonstrated excellent results for hard-to-understand words, with a rate of 92.48% (Table 8), which illustrates the tool's strength in tackling vocabulary learning barriers. Most learners in the

experimental group reported being happy with their experience using AI tools.

The intervention received overwhelmingly positive feedback from participants, with response rates between 94.23% and 97.45% (Table 9) reflecting their perception of its usefulness, together with their enjoyment while using it.

Table 8: Retention across vocabulary difficulty levels (Performance by word category)

Word category	Group	Immediate post-test (M)	SD	Delayed post-test (M)	SD	Retention rate (%)
Easy words	Experimental group	88.23	4.21	85.67	5.01	97.10
	Control group	82.45	5.34	77.12	6.12	93.55
Moderate words	Experimental group	78.56	6.34	73.45	7.21	93.50
	Control group	69.45	7.89	63.45	8.23	91.34
Difficult words	Experimental group	65.12	8.11	60.23	8.97	92.48
	Control group	56.23	8.89	50.12	9.45	89.10

Table 9: Student satisfaction with AI tools

Survey item	Mean (M)	SD	Percentage of 4 or 5 responses (%)
Ease of use of AI tools	4.78	0.42	96.67
Relevance of AI-generated exercises	4.62	0.51	94.23
Impact of AI feedback on learning	4.80	0.36	97.45
Enjoyment of the learning experience	4.75	0.49	95.89

Self-reported questionnaire results; 5-point Likert scale: 1 = Strongly disagree, 5 = Strongly agree

Table 10 shows that there is a strong positive correlation between engagement levels and retention scores, indicating that increased engagement facilitated by AI tools contributed to better vocabulary retention. The experimental participants dedicated more time to their assignments, yielding test results showing a t-value

of 8.56 and a p-value of 0.000 (Table 11). The time-on-task difference between control and experimental groups reached statistical significance based on t-values of 5.45 and p-values of 0.000. Students in the experimental group spent additional time working on their assignments, which fundamentally contributed to superior vocabulary retention.

Table 10: Correlation analysis between engagement and retention

Variable 1	Variable 2	Correlation coefficient (r)	P-value
Engagement levels	Immediate post-test scores	0.72	0.000
Engagement levels	Delayed post-test scores	0.68	0.000

Table 11: Time-on-task analysis (in minutes per session)

Group	Mean (M)	SD	T-value	P-value
Experimental group	35.42	3.21	8.56	0.000
Control group	27.34	4.12	5.45	0.000

According to Table 12, the experimental AI intervention proved equally successful in terms of results between male and female participants in the immediate post-testing and delayed post-testing.

Table 12: Gender-based differences in retention (experimental group)

Test	Gender	Mean (M)	SD	T-value	P-value
Immediate post-test	Male	78.12	8.21	1.23	0.219
	Female	79.23	7.89		
Delayed post-test	Male	74.56	8.34	1.11	0.267
	Female	75.78	8.01		

Research findings from this investigation show improved comprehension regarding AI effects on educational language achievement, specifically observed in Saudi learning environments. Study results between experimental and control

participants allowed assessment of AI technology effects on terminology learning processes, student interest, and memory improvement, and satisfaction. The obtained outcomes have value beyond language instruction advancement because they enable critical

evaluation of how technology affects educational results.

4.1. Engagement and time-on-task

The experimental participants who used AI technology tools achieved higher engagement levels, together with longer periods of classroom participation, in comparison to the control participants. Studies by [Sajja et al. \(2024\)](#) and others have documented that AI educational tools boost learner engagement through their customizable learning interface. Student engagement increases while vocabulary tasks extend, thanks to AI tools that adapt learning environments based on student progression and dynamically provide real-time feedback.

The rise in both time-on-task and engagement among the experimental group is directly related to the design elements of the AI tool being 'gamified' and its ability to be differentiated to suit each learner individually. Through the lens, transforming learning a specific set of vocabulary words into a 'gaming' activity with instant gratification and visible milestones ([Alenezi, 2023](#)), the effect of the AI tool promoted a sense of achievement and fostered sustained intrinsic motivation among learners. Unlike the fixed exercises available in textbooks, the dynamic adaptation of content and tasks according to the learners' progress stimulated a 'personalized learning pathway' and thereby increased learner autonomy and a 'sense of ownership' in learning, as advocated in a study showing a correlation with increased 'cognitive processing' and learning 'retention' ([Huang et al., 2023](#)).

However, while increased time-on-task is a desirable attribute, it is important for administrators and educators to be mindful of the quality and level of cognitive engagement invested during the increased period. A lot of dedication towards menial tasks, regardless of whether they take a gamification angle, may lack significant learning value and critical thinking abilities if these tasks do not present adequate challenges towards higher-order thinking. There is a need for further research aimed at qualifying levels of engagement, whether passive and active engagements, within problem-solving, so that a profound level of linguistic understanding is promoted rather than simply the activity itself being completed satisfactorily.

Research must evaluate the impact of Saudi cultural factors on how learners utilize AI-powered teaching instruments. The study focuses on learner engagement, but the participants' cultural origin influences their technological reactions and understanding ([O'Brien and Toms, 2008](#)). Digital education in Saudi Arabia is attracting rising attention, although technological accessibility and digital competency differ across economic and geographic areas. The research omitted variables that could affect participants' level of engagement in the entire process. These results cannot be

considered universally applicable to all Saudi Arabian learners.

4.2. Retention rates

The greatly increased retention of vocabulary noted among the experimental group can be firmly linked with the sophisticated use of spaced repetition algorithms by the AI tool. The spaced repetition algorithms used in the tool allowed for a calibrated level of reviewing based on the specific forgetting curves of each learner, so that the recalling of words took place when the probability of memory loss was greatest, thus helping to strengthen the neuronal pathways connected with long-term memorization of words. Simultaneously, the adaptive learning function allowed for instant corrective learning and dynamic adjustment of exercise levels in real-time, which directly targeted learning gaps and prevented the solidification of errors becoming 'fossilized.' This dynamic targeting is considerably different from the one-size-fits-all approach used by standard learning platforms, which do not separately factor the variabilities of learning among learners and the peculiar requirements related to learning words with differing characteristics. The dynamic possibility of content delivery offered by the AI tool greatly increased the learning efficiency and simultaneously allowed learners' cognizance with words to be deeply grounded beyond the level of simple memorization ([Lindsey et al., 2014](#); [Halkiopoulou and Gkintoni, 2024](#)).

The major benefit of AI-based tools exists in their adaptive capability to respond to students' learning progress through specially tailored assistance strategies during problematic items. According to [Halkiopoulou and Gkintoni \(2024\)](#), the usage of AI for individualized feedback improves vocabulary retention through personalized learning approaches. The experimental group utilized AI tools that detected learners' knowledge gaps so they could adjust educational difficulty settings. The system delivered better vocabulary learning outcomes with superior retention results, particularly for challenging words that learners normally would fail to learn.

The expected increase in learning retention from short-term research needs further validation for its durability as long-term language proficiency. The measured intervention results delivered significant retention outcomes immediately, without understanding how these findings will persist in the long-term vocabulary retention ([Duncan et al., 2020](#); [Abdulbaki et al., 2025](#)). According to [Ruiz-Rojas et al. \(2024\)](#), the purpose of AI tools involves maximizing immediate learning success, yet their persistent influence lacks clarity. Future research should analyze what happens to vocabulary acquired using AI tools through periodic retention checks performed in multiple months following the intervention.

AI systems undermine the vital components of language learning, which depend upon social and contextual experiences needed for long-term vocabulary retention in actual situations. The learning process extends beyond vocabulary memorization because it involves genuine social interaction and cultural understanding, together with practical language application in multiple settings. Understanding vocabulary with AI solutions becomes possible, but authentic exchanges coupled with experiential learning produce a better range of language mastery. Higher retention discovered in this research, however, does not necessarily mean students will attain superior language proficiency.

4.3. Satisfaction and user experience

Students who used AI-powered tools in the experimental segment displayed significantly stronger satisfaction levels than their counterparts in the control segment, while showing positive attitudes about both tool usability and interactive capabilities. Research has shown that students show strong satisfaction with AI tools because these tools customize learning interactions alongside immediate feedback provision (Huang et al., 2023). This study demonstrates that AI successfully improves learner satisfaction because it delivers intuitive interfaces that improve student engagement and maintain their continued interactions.

The success of educational tools depends highly on satisfaction, but satisfaction alone fails to produce enhanced learning results (Kirkwood and Price, 2014). Student satisfaction scores are likely to increase due to new factors in the AI tools, though such easy approval does not ensure deeper educational gains. Research data reveals that students who feel satisfied with technological learning do not necessarily demonstrate enhanced academic results (Zheng and Xiao, 2024). Studies need to investigate better how learner satisfaction rates reflect real learning achievements, specifically in AI-assisted language teaching.

The level of satisfaction with AI tools in Saudi Arabia could be affected by cultural features that exist only in this country (Albalwi, 2023). The rise of educational technology throughout Saudi Arabia has triggered expanded use of AI tools, but students show different familiarity levels with these tools depending on whether they come from urban or rural settings (Aloufi et al., 2024). The satisfactory experience of AI tools demonstrated in this study merits research on how learners from diverse educational backgrounds and regions across Saudi Arabia respond to these AI tools. The research findings can establish the nationwide applicability of the obtained positive outcomes.

4.4. Gender equality in learning outcomes

The experimental group revealed a matching lack of meaningfully different engagement and retention patterns between genders. The finding stands out in

Saudi Arabia, given how gender norms frequently shape educational experiences. Students perform higher in language-based learning among females than males across educational environments, especially when learning vocabulary. AI tools appear to create an equivalent learning environment for vocabulary retention because both experimental and control group students, regardless of their gender, achieved similar vocabulary results.

The research by Miao and Holmes (2021) supported this finding by showing how AI enables access to customized educational paths, which help create equal educational opportunities across genders. AI tools can serve individual learner needs across genders through their approach of personalized feedback alongside adaptive academic paths. The analysis suggests that AI tools enhanced learning extensively because both male and female participants across the group could equally advance in their studies after their interaction with the system.

Overall, these promising research findings require moderate interpretation because of their potential risks. The influence of gender dynamics on education extends past technology usage, complicating these relationships (Rudman et al., 2021). Social expectations, together with cultural norms, may have been contributing elements that shaped male and female students' use of AI tools. Research needs to investigate how AI tools engage with educational gender dynamics while recognizing cultural differences and their impact on educational outcomes.

4.5. The role of AI in language learning

The outcome of this study adds new evidence to the expanding investigation of AI functions in language acquisition. Research outcomes demonstrate that AI technology delivers substantial improvements in student engagement, combined with increased retention levels and better satisfaction when implemented for customized educational content delivery and immediate feedback mechanisms. The results show that AI tools generate superior educational outcomes when they are properly integrated throughout educational systems.

Students should not see AI tools as replacements for standard teaching practices, yet they benefit considerably from their implementation to boost educational experiences. According to Garrison and Kanuka (2004), technology should help enhance typical classroom interactions instead of replacing them. AI tools succeeded in maintaining student interest while supporting academic persistence during this research, yet students couldn't achieve complete social and cultural educational benefits, which only emerged through traditional teacher-student relationships in classroom learning.

AI tools should not substitute complete educational methodologies that promote profound linguistic understanding; however, they effectively

enhance traditional approaches for language learning. Numerical findings from this research highlight the critical necessity of examining how educational institutions leverage AI technology. AI technology delivers tailored learning experiences, yet creates problems regarding system accessibility for all students, and it affects both student equity and comprehensive long-term assessment results. Dimitriadou and Lanitis (2023) explained that AI devices obtain their best performance when properly integrated into proper teaching settings. The research delivers critical findings about AI vocabulary learning potential while arguing for additional intensive research into its extensive applications over time.

5. Conclusion and recommendations

The study evaluated the effects of AI tools on vocabulary retention of Saudi Arabian foreign language students by demonstrating the potential technology holds to improve language education results. AI-powered learning tools brought about better retention outcomes and increased engagement, together with higher satisfaction rates for their users.

The students who used AI tools demonstrated superior vocabulary retention while conducting extended tasks with higher engagement levels when compared to students from the control group. AI tools demonstrate their effectiveness as language learning aids because they deliver tailored educational resources that react to individual needs while achieving better educational results. Future implementation of AI in education must consider important aspects that are discussed in this study. The promising results from this research study raise important questions about how long-term effects from AI vocabulary learning systems impact students' overall mastery of language. The study insufficiently investigated the various cultural factors that impact student interaction with AI tools through analysis of digital literacy competencies and contrasting technological capabilities in different regions. The positive reception of AI tools in this study presents important challenges for achieving equitable access to these tools across all Saudi demographic groups in the Saudi educational context.

To fill these research gaps, scientists should examine how permanent AI can maintain vocabulary knowledge while investigating retention effects on varied groups of language learners. Additional research should evaluate the complete educational advantages of AI language curriculum inclusion, together with diverse assessment methods of how AI can enhance educational processes.

This study investigates AI techniques in education by delivering important findings that benefit educational stakeholders and researchers focused on technological language education advancement.

Funding

This research is supported by King Khalid University (Grant number: RGP2/681/46).

Acknowledgment

The authors extend their appreciation to the Deanship of Research and Graduate Studies at King Khalid University for funding this work through the Large Research Project under grant number RGP2/681/46.

Compliance with ethical standards

Ethical considerations

This study was approved by the Ethical Committee of the Deanship of Scientific Research at King Khalid University, Saudi Arabia (Ref. No. RGP2/681/46). All participants provided informed consent and were informed of their right to withdraw at any time. Data were collected anonymously, kept confidential, and used solely for research purposes, with appropriate measures taken to ensure data privacy in the use of AI-based tools.

Conflict of interest

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

- Abdulbaki S, Harrathi H, Khasawneh MA, and Tashtoush MA (2025). A comparative study of national, gender, and academic differences in smartphone addiction among students from Jordan, Saudi Arabia, Oman, and the UAE. *International Journal of Advanced and Applied Sciences*, 12(8): 101-109. <https://doi.org/10.21833/ijaas.2025.08.010>
- Albalwi ASM (2023). Investigating the influence of learning-centred leadership on authentic instruction and the contribution to enhancing the learning culture: A case study of four Saudi Arabian primary schools. Ph.D. Dissertation, Newcastle University, Newcastle, UK.
- Aldowsari BI and Aljebreen SG (2024). The impact of using a ChatGPT-based application to enhance Saudi students' EFL vocabulary learning. *International Journal of Language and Literary Studies*, 6(4): 380-397. <https://doi.org/10.36892/ijlls.v6i4.1955>
- Alenezi A (2023). Teacher perspectives on AI-driven gamification: Impact on student motivation, engagement, and learning outcomes. *Information Technologies and Learning Tools*, 97(5): 138-148. <https://doi.org/10.33407/itlt.v97i5.5437>
- Al-Faqih MM (2015). English grammar achievement level among secondary school leavers in Saudi Arabia. Ph.D. Dissertation, Leeds Beckett University, Leeds, UK.
- Aloufi F, Tashtoush MA, Shirawia N, Tashtoush R, and Az-Zo'bi E (2024). Internet of Things in education: Teachers' perspectives, practices and challenges. *WSEAS Transactions on Computer Research*, 12: 429-442. <https://doi.org/10.37394/232018.2024.12.42>
- Beirat MA, Tashtoush DM, Khasawneh MA, Az-Zo'bi EA, and Tashtoush MA (2025). The effect of artificial intelligence on

- enhancing education quality and reduce the levels of future anxiety among Jordanian teachers. *Applied Mathematics and Information Sciences*, 19(2): 279-290.
<https://doi.org/10.18576/amis/190205>
- Chen YL, Hsu CC, Lin CY, and Hsu HH (2022). Robot-assisted language learning: Integrating artificial intelligence and virtual reality into English tour guide practice. *Education Sciences*, 12(7): 437.
<https://doi.org/10.3390/educsci12070437>
- Dimitriadou E and Lanitis A (2023). A critical evaluation, challenges, and future perspectives of using artificial intelligence and emerging technologies in smart classrooms. *Smart Learning Environments*, 10: 12.
<https://doi.org/10.1186/s40561-023-00231-3>
PMid:40478141 PMCID:PMC9900563
- Duncan EM, Bennett T, and Gillies K (2020). Assessing effective interventions to improve trial retention: Do they contain behaviour change techniques? *Trials*, 21: 213.
<https://doi.org/10.1186/s13063-020-4151-4>
PMid:32085717 PMCID:PMC7035706
- Farmer TW, Bierman KL, Hall CM, Brooks DS, and Lee DL (2021). Tiered systems of adaptive supports and the individualization of intervention: Merging developmental cascades and correlated constraints perspectives. *Journal of Emotional and Behavioral Disorders*, 29(1): 3-13.
<https://doi.org/10.1177/1063426620957651>
- Fu J (2009). A study of learning styles, teaching styles and vocabulary teaching strategies in Chinese primary school: How do they differ and how can they be integrated? Available online at:
<http://www.diva-portal.org/smash/get/diva2:228945/fulltext01>
- Garrison DR and Kanuka H (2004). Blended learning: Uncovering its transformative potential in higher education. *The Internet and Higher Education*, 7(2): 95-105.
<https://doi.org/10.1016/j.iheduc.2004.02.001>
PMCID:PMC8236236
- Halabi M (2017). The role of the tutors in supporting learners in a higher education distance language learning programme environment in Saudi Arabia. Ph.D. Dissertation, University of Sheffield, Sheffield, UK.
- Halkiopoulou C and Gkintoni E (2024). Leveraging AI in e-learning: Personalized learning and adaptive assessment through cognitive neuropsychology—A systematic analysis. *Electronics*, 13(18): 3762.
<https://doi.org/10.3390/electronics13183762>
- Hatamleh HA, Alsaadi O, Alkhafaji B, Khasawneh MA, and Tashtoush MA (2025). Game-based and AI-driven engagement strategies to combat demotivation in foreign language learning. *International Journal of Advanced and Applied Sciences*, 12(3): 119-130.
<https://doi.org/10.21833/ijaas.2025.03.013>
- Helsing D, Howell A, Kegan R, and Lahey L (2008). Putting the "development" in professional development: Understanding and overturning educational leaders' immunities to change. *Harvard Educational Review*, 78(3): 437-465.
<https://doi.org/10.17763/haer.78.3.8881759g1qm54660>
- Huang AY, Lu OH, and Yang SJ (2023). Effects of artificial intelligence-enabled personalized recommendations on learners' learning engagement, motivation, and outcomes in a flipped classroom. *Computers & Education*, 194: 104684.
<https://doi.org/10.1016/j.compedu.2022.104684>
- Kirkwood A and Price L (2014). Technology-enhanced learning and teaching in higher education: what is 'enhanced' and how do we know? A critical literature review. *Learning, Media and Technology*, 39(1): 6-36.
<https://doi.org/10.1080/17439884.2013.770404>
- Lindsey RV, Shroyer JD, Pashler H, and Mozer MC (2014). Improving students' long-term knowledge retention through personalized review. *Psychological Science*, 25(3): 639-647.
<https://doi.org/10.1177/0956797613504302>
PMid:24444515
- Miao F and Holmes W (2021). AI and education: A guidance for policymakers. UNESCO Publishing, Paris, France.
- O'Brien HL and Toms EG (2008). What is user engagement? A conceptual framework for defining user engagement with technology. *Journal of the American Society for Information Science and Technology*, 59(6): 938-955.
<https://doi.org/10.1002/asi.20801>
- Rudman LA, Glick P, and Glick PS (2021). The social psychology of gender: How power and intimacy shape gender relations. Guilford Publications, New York, USA.
- Ruiz-Rojas LI, Salvador-Ullauri L, and Acosta-Vargas P (2024). Collaborative working and critical thinking: Adoption of generative artificial intelligence tools in higher education. *Sustainability*, 16(13): 5367.
<https://doi.org/10.3390/su16135367>
- Sajja R, Sermet Y, and Demir I (2025). End-to-end deployment of the educational AI hub for personalized learning and engagement: A case study on environmental science education. *IEEE Access*, 13: 55169-55186.
<https://doi.org/10.1109/ACCESS.2025.3554222>
- Sajja R, Sermet Y, Cikmaz M, Cwiertny D, and Demir I (2024). Artificial intelligence-enabled intelligent assistant for personalized and adaptive learning in higher education. *Information*, 15(10): 596.
<https://doi.org/10.3390/info15100596>
- Srinivasa KG, Kurni M, and Saritha K (2022). Harnessing the power of AI to education. In: Srinivasa KG, Kurni M, and Saritha K (2022). *Learning, teaching, and assessment methods for contemporary learners: 311-342*. Springer, Singapore, Singapore. https://doi.org/10.1007/978-981-19-6734-4_13
- Teng MF and Reynolds BL (2024). Researching incidental vocabulary learning in a second language. Routledge, London, UK. <https://doi.org/10.4324/9781003270782-1>
- Zheng Y and Xiao A (2024). A structural equation model of online learning: Investigating self-efficacy, informal digital learning, self-regulated learning, and course satisfaction. *Frontiers in Psychology*, 14: 1276266.
<https://doi.org/10.3389/fpsyg.2023.1276266>
PMid:38274678 PMCID:PMC10809156