

Comparative Study of AI-based Applications and Traditional Methods on Vocabulary and Spelling Proficiency: Implications for Educational Recommendation

Geraldine L. Gaddi¹, Virginia S. Ariza²

¹Holy Infant College of Tacloban City, Inc., Tacloban City, Leyte, Philippines

²Eastern Visayas State University Carigara Campus

DOI: <https://doi.org/10.51584/IJRIAS.2026.11080150>

Received: 3 September 2026; Accepted: 5 September 2026; Published: 19 September 2026

ABSTRACT

This quasi-experimental study examined whether AI-based applications were more effective than traditional teaching methods in improving vocabulary and spelling proficiency among Grade 9 English language learners in Basey II District, Philippines. The study addressed the need to determine whether AI-based learning tools provide better learning outcomes than commonly used teaching methods in developing learners' language skills. Thirty learners from intact sections participated in the study and were compared using a pretest-posttest design, with 15 learners exposed to AI-based applications and 15 taught through traditional methods. A researcher-made test measured vocabulary and spelling proficiency before and after the intervention. The pretest results showed lower vocabulary performance in the traditional group (MPS = 29.7) than in the AI-based group (MPS = 40.7). In contrast, spelling performance was similar between the traditional group (MPS = 50.0) and the AI-based group (MPS = 53.0). Both groups showed large improvements after the intervention. Vocabulary MPS increased to 92.33 in the traditional group and 93.33 in the AI-based group, while spelling MPS increased to 92.67 and 94.33, respectively. Paired-samples t-tests showed significant improvements in vocabulary and spelling scores within both groups ($p < .001$). Although a significant difference was found in pretest vocabulary scores ($p = .013$), no significant difference was found between the groups in posttest vocabulary ($p = .844$) or spelling ($p = .692$). These findings indicate that both AI-based and traditional teaching methods improved vocabulary and spelling proficiency. However, the AI-based approach did not produce significantly higher posttest performance than the traditional approach. The findings suggest that AI-based applications may serve as a useful supplementary teaching tool, while traditional methods remain effective, particularly in settings with limited access to technology.

Keywords: AI applications, literacy enhancement, spelling proficiency, traditional methods, vocabulary proficiency.

INTRODUCTION

Artificial Intelligence (AI) has emerged as one of the most powerful drivers of change in the 21st century, fundamentally redefining how individuals work, communicate, and learn. AI refers to computer systems that can perform tasks that typically require human intelligence, such as language processing, decision-making, and adaptive learning. However, integrating AI into education presents both opportunities and challenges. Globally, AI-powered platforms provide real-time feedback and personalized learning pathways that adapt to individual student needs—advantages often limited in traditional classrooms due to size and resource constraints (Mekheimer, 2025). However, rural schools such as those in the Basey II District face barriers that hinder AI's full potential. These include limited digital infrastructure, unequal access to devices, and insufficient teacher training, which collectively widen the educational gap between urban and rural learners, particularly in critical foundational skills like vocabulary and spelling (Rosqueta, 2025).

The researchers firsthand encounter with these issues highlighted the necessity of structured, context-sensitive recommendations for literacy enhancement that thoughtfully combine AI-based tools with traditional instructional methods. This integration ensured that technology supports, rather than replaces, the development of genuine vocabulary and spelling proficiency. Within this context, Grade 9 learners in Basey II District continue to face persistent difficulties in vocabulary and spelling proficiency - essential for academic success and effective communication. As AI becomes increasingly embedded in global education systems, it is critical to explore how to adapt these tools to local realities. The researcher's experiences of both the potential and limitations of AI in the classroom reinforce the need for a literacy program that maximizes AI's benefits while addressing practical challenges. This study aimed to compare the effectiveness of AI-based applications and traditional teaching approaches on vocabulary and spelling proficiency among Grade 9 learners in Basey II District.

METHODOLOGY

Research Design

This study used a quasi-experimental research design, commonly used to evaluate the effectiveness of educational interventions when random assignment of participants is not feasible. The design allowed the researcher to compare AI-based approaches with traditional instructional approaches in terms of vocabulary and spelling proficiency among Grade 9 English language learners. Specifically, the study involved two groups of learners: an AI-based group that used AI-based applications as part of their English language instruction, and a traditional group that followed conventional teaching methods, such as teacher-led instruction, written tasks, dictation exercises, board work, use of worksheets and flashcards. Both groups underwent pretest assessments to establish baseline proficiency levels in vocabulary and spelling. After the intervention period, posttest assessments were conducted to measure any improvements or changes in proficiency. A quasi-experimental design was particularly appropriate for this study as it enabled a comparative evaluation of the instructional approaches (AI-based and traditional) and their corresponding learning outcomes, while accommodating real-world classroom conditions in which full randomization may not be feasible.

Research Locale

The study was conducted in Basey II District in Basey, which includes thirty (30) elementary schools and three (3) secondary schools under the leadership of Dr. Reldin R. Cajés, District In-Charge. The study respondents were thirty (30) Grade 9 learners from Old San Agustin National High School. The district was selected because it provides an appropriate setting to examine vocabulary and spelling instruction among secondary learners in a real classroom environment. Basey is situated across the San Juanico Strait from Tacloban City and is known for its accessible educational institutions and diverse learning contexts. Among the district's secondary schools, Old San Agustin National High School, located in Barangay Old San Agustin, served as the research locale. The school is located east of the town proper and is about a 10–15 minute tricycle or motorcycle ride away. The school was chosen as the research locale because it is accessible to the researcher and suitable for a quasi-experimental study, making it an appropriate venue for examining the effectiveness of AI-based and traditional instructional approaches in improving learners' vocabulary and spelling proficiency.

Research Respondents

The respondents in this study were 30 Grade 9 English language learners officially enrolled at Old San Agustin National High School in the Basey II District. These students were selected because they were at a critical stage in their English language learning, particularly in developing vocabulary and spelling proficiency. The study used intact sections as clusters for group assignment, with one section serving as the AI-based group exposed to AI-based applications and another serving as the traditional group following traditional instructional methods. Using intact classes minimizes disruption to school schedules, supports ethical considerations, and improves the feasibility of implementing the intervention. To ensure comparability between the groups, the researchers administered a pretest on vocabulary and spelling proficiency to all sections. They assigned sections with similar baseline scores to the AI-based and traditional groups. This approach allows for practical implementation of the intervention while maintaining fairness in group assignment. The total number of respondents is $N = 30$,

representing all students in the selected classes. The sample size is sufficient for statistical analysis, including t-tests to compare pretest and posttest scores. The study considered respondents' access to technology; regular attendance, parental consent, and student willingness to participate were secured to ensure ethical compliance and active involvement throughout the intervention period.

Research Instrument

The study used researcher-developed pretest and posttest instruments to assess Grade 9 learners' vocabulary and spelling proficiency and compare AI-based and traditional methods. The 40-item test included 20 multiple-choice vocabulary items covering word meanings, synonyms, antonyms, and contextual usage, and 20 spelling items requiring learners to rewrite or type misspelled words correctly in sentences. Each item was worth one point. The traditional group answered printed questionnaires with a pen. In contrast, the AI-based group completed the same test through a web-based platform accessed via a quiz link or QR code on an internet-connected device. The AI-based test provided immediate evaluation, total scores, and real-time results accessible to the researcher. It was available in a plain text format and a gamified format with a colorful interface. Both groups answered identical test items using different modes of administration. The researchers validated the instruments based on the DepEd curriculum (DepEd, 2020) and recognized spelling standards (Merriam-Webster, 2023) to ensure relevance and suitability for Grade 9 learners. English teaching and curriculum experts reviewed the test items to determine their alignment with the Grade 9 Most Essential Learning Competencies (MELCs), accuracy, clarity, appropriateness, and difficulty level, and their suggestions were used to improve the instruments. The researchers also tried the tests with a small group of Grade 9 learners not included in the main study to assess instruction clarity, difficulty, completion time, and scoring. The researchers used the trial results to revise unclear items and improve the instruments. The researchers also conducted live testing of the AI-based questionnaires to assess reliability and scorability. Through these validation and testing procedures, the instruments were established as appropriate for measuring vocabulary and spelling proficiency.

Data Gathering Procedure

The data collection process consisted of four phases: preparation, pre-implementation, implementation, and post-implementation. During the preparation phase, the researcher secured approval from the Dean of the Graduate School and obtained permission from the Schools Division Superintendent, District Supervisor, and School Head of Old San Agustin National High School. The researcher also obtained written consent from the parents or guardians of the selected Grade 9 learners, emphasizing voluntary participation, confidentiality, and anonymity. The researchers prepared the research instruments, validated them with experts, and pilot-tested them to ensure clarity, appropriateness, and reliability. During the pre-implementation phase, participants were identified based on the study criteria and grouped into AI-based and traditional groups, considering their ability and background. The researchers oriented learners to the study's purpose, procedures, duration, expectations, and ethical considerations. The researchers then administered a pretest to both groups to determine baseline vocabulary and spelling proficiency. The researchers recorded, organized, and tabulated the scores to establish the learners' initial performance and provide a basis for comparison with the posttest results.

The implementation phase lasted 10 days, with the same teacher facilitating identical lessons for both groups within the same timeframe. The AI-based group used interactive and gamified activities through digital flashcards, quizzes, spelling games, word searches, sentence completion, and AI-assisted platforms such as ChatGPT, Quiz Maker, Qzzr.com, and Wayground. Meanwhile, the traditional group used teacher-led activities such as flashcards, dictation, worksheets, board work, sentence completion, matching exercises, and written spelling tasks. Both approaches provided structured activities that gradually developed vocabulary and spelling skills. AI-based activities provided immediate feedback, interactive practice, and automatic score monitoring, while traditional activities emphasized handwriting, repetition, teacher feedback, peer interaction, and classroom engagement. The AI-based group's progress was monitored through automatically recorded scores, while the traditional group's results were manually checked, recorded, and tabulated. During the post-implementation phase, the researcher administered a posttest to both groups to measure changes in vocabulary and spelling proficiency. The researcher statistically analyzed the pretest and posttest scores to determine the effectiveness of the two teaching approaches and compare their results. The researcher also recorded observations and teacher notes on learner engagement, participation, interaction with AI tools, peer collaboration, and responses to

feedback. These observations complemented the quantitative results and provided additional information about the learning process. The findings then formed the basis for recommendations to improve instructional strategies and consider using AI tools alongside traditional teaching methods to address different learner needs.

Data Analysis Procedure

The data were analyzed using descriptive and inferential statistics. Frequency, percentage, mean percentage score (MPS), and standard deviation were used to summarize learners' performance. Paired-samples t-tests were used to examine pretest-to-posttest changes within each instructional group. Independent-samples t-tests were used to describe the unadjusted between-group differences. Because the pretest vocabulary scores differed significantly between groups, ANCOVA was additionally conducted using posttest vocabulary as the dependent variable, instructional approach as the fixed factor, and pretest vocabulary as the covariate. A corresponding ANCOVA was conducted for spelling using pretest spelling as the covariate. The ANCOVA models were used to statistically control for baseline proficiency and provide an adjusted estimate of the between-group effect. The assumptions of ANCOVA, including linearity, homogeneity of regression slopes, normality of residuals, and homogeneity of variance, were considered. Statistical significance was evaluated at $\alpha = .05$. Cohen's d was reported alongside p-values to describe the practical magnitude of the observed differences, while partial eta squared was reported for the ANCOVA group effects.

Ethical Considerations

This study adhered to strict ethical standards to protect participants' rights, welfare, and dignity. The researchers obtained parental consent and student assent, and participants were fully informed of the study's purpose, procedures, benefits, and right to withdraw. The researcher ensured confidentiality and anonymity by coding personal information and keeping it accessible only to the researcher. The AI-based applications were educational, age-appropriate, and safe, with no harm, and the researcher provided guidance as needed. The researcher secured permissions from school authorities and relevant offices and maintained honesty and integrity throughout data collection, analysis, and reporting.

RESULTS AND DISCUSSION

Pretest and Posttest Vocabulary Proficiency Using Traditional and AI-Based Approaches

Table 1 presents respondents' spelling proficiency levels in the pretest and posttest under the traditional and AI-based approaches. It shows the distribution of learners by mastery level, including frequency, percentage, Mean Percentage Score (MPS), qualitative description, and standard deviation before and after the intervention. This presentation compares respondents' spelling proficiency levels before and after implementing the two instructional approaches and determines the extent of improvement in each group. The traditional approach substantially improved vocabulary proficiency. In the pretest, 9 out of 15 respondents (60.00%) were at the Low Mastery level, and 6 (40.00%) were at the Average Mastery level, with none reaching the higher mastery levels. The pretest mean score was 29.67, interpreted as Low Mastery, while the standard deviation of 9.72 indicated relatively similar performance among respondents. After the intervention, 9 respondents (60.00%) reached the Mastered level, 3 (20.00%) reached Closely Approximating Mastery (CAM), 2 (13.30%) reached Moving Towards Mastery (MTM), and 1 (6.70%) remained at Average Mastery. The posttest mean increased to 92.33, interpreted as CAM, showing substantial improvement that may be associated with structured instruction, repetition, drills, and reinforcement.

Table 1. Pretest and Posttest Vocabulary Proficiency Using Traditional and AI-Based Approaches

Level of Mastery		Traditional Approach (n = 15)				AI-based Approach (n= 15)			
MPS	Qualitative Description	Pretest		Posttest		Pretest		Posttest	
		<i>f</i>	%	<i>f</i>	%	<i>f</i>	%	<i>f</i>	%
96 – 100	Mastered (M)	0	0.00	9	60.00	0	0	7	46.70
86 – 95	Closely Approximating Mastery (CAM)	0	0.00	3	20.00	0	0	5	33.30

66 – 85	Moving towards Mastery (MTM)	0	0.00	2	13.30	1	6.70	3	20.00
35 – 65	Average Mastery (AM)	6	40.00	1	6.70	11	73.30	0	0.00
15 – 34	Low Mastery (LM)	9	60.00	0	0.00	3	20.00	0	0.00
Overall Mean		29.67		92.33		40.67		93.33	
Qualitative Description		LM		CAM		AM		CAM	
SD		9.72		16.99		12.80		9.57	
Note. f=frequency; SD=standard deviation; MPS=Mean percentage score									

This finding supports Skinner’s Behaviorist Theory, which emphasizes habit formation, repetition, imitation, and reinforcement in language learning. The improvement in the traditional group may also reflect Piaget’s Cognitive Theory, which emphasizes the role of cognitive development in language acquisition. Meaningful classroom activities, discussions, and exercises may have helped learners process and apply vocabulary more effectively. The findings also support Reed et al. (2016), who emphasized that explicit instruction, repeated practice, and memory reinforcement contribute to vocabulary and spelling acquisition. However, the higher posttest standard deviation of 16.99 indicates greater variation in learners’ performance after the intervention. Differences in local living conditions, travel distance to school, and family support due to labor responsibilities may have affected the time and assistance available for learning. Despite these differences, most learners achieved mastery or near-mastery levels.

The AI-based approach likewise produced substantial improvement in vocabulary proficiency. Before the intervention, 11 respondents (73.30%) were at Average Mastery, 3 (20.00%) were at Low Mastery, and 1 (6.70%) was at Moving Towards Mastery, with none reaching CAM or Mastered. The pretest mean of 40.67 indicated Average Mastery, which was higher than the traditional group’s initial mean, while the standard deviation of 12.80 showed moderate variation in performance. After the intervention, 7 respondents (46.70%) reached Mastered, 5 (33.30%) reached CAM, and 3 (20.00%) reached MTM, with none remaining at Average or Low Mastery. The posttest mean increased to 93.33, interpreted as CAM. This improvement may be linked to the interactive, adaptive, and engaging nature of AI-based activities, including immediate feedback, personalized practice, and repeated exposure. The findings for the AI-based group support Krashen’s Input Hypothesis, which suggests that language acquisition is promoted through comprehensible, meaningful input slightly beyond learners’ current proficiency. They also support Krashen’s Affective Filter Hypothesis, as interactive and gamified AI activities may create a less intimidating learning environment and encourage participation. The results are also consistent with Cummins’ theory of Cognitive Academic Language Proficiency (CALP), which emphasizes contextualized vocabulary development for academic learning. In addition, adaptive learning technologies can adjust activities based on learner performance, allowing learners to practice at their own pace while receiving immediate feedback. The lower posttest standard deviation of 9.57 suggests more consistent performance among the AI-based respondents after the intervention.

The comparative results show that both groups improved substantially from pretest to posttest. The traditional group increased from an MPS of 29.67 to 92.33, while the AI-based group increased from 40.67 to 93.33. The AI-based group had a slightly higher posttest mean and more consistent gains overall, while the traditional group had a higher proportion of learners reaching the Mastered level. These findings are consistent with Torres (2025), Liu (2025), Zhang et al. (2022–2025), and Rahman (2025), who reported that AI-supported learning can improve vocabulary acquisition, retention, motivation, engagement, and self-directed learning. However, the findings also recognize the continued effectiveness of traditional approaches and the concerns raised by Bao (2025) and Khalifeh et al. (2026) regarding possible overreliance on AI, reduced deeper cognitive processing, and a false sense of mastery. Overall, the findings indicate that both traditional and AI-based approaches improved vocabulary proficiency. Although the AI-based group obtained a slightly higher posttest MPS of 93.33 than the traditional group’s 92.33, the difference was small, suggesting that AI should be viewed as a supportive instructional tool rather than a replacement for effective teaching. The traditional group’s strong results further demonstrate the value of structured instruction, practice, and teacher guidance. Therefore, integrating AI-supported activities with established teaching strategies may provide a useful approach for addressing learners’ vocabulary needs and supporting intervention programs such as ARAL and SARP.

Pretest and Posttest Spelling Proficiency Using Traditional and AI-Based Approaches

Table 2 presents respondents' spelling proficiency levels in the pretest and posttest under the traditional and AI-based approaches. It shows the distribution of learners by mastery level, including frequency, percentage, Mean Percentage Score (MPS), qualitative description, and standard deviation before and after the intervention. This presentation compares respondents' spelling proficiency levels before and after implementing the two instructional approaches and determines the extent of improvement in each group.

Table 2. Pretest and Posttest Spelling Proficiency Using Traditional and AI-Based Approaches

Level of Mastery		Traditional Approach (n = 15)				AI-based Approach (n= 15)			
MPS	Qualitative Description	Pretest		Posttest		Pretest		Posttest	
		<i>f</i>	%	<i>f</i>	%	<i>f</i>	%	<i>f</i>	%
96 – 100	Mastered (M)	0	0.00	7	46.67	0	0.00	9	60.00
86 – 95	Closely Approximating Mastery (CAM)	0	0.00	4	26.67	1	6.67	3	20.00
66 – 85	Moving towards Mastery (MTM)	3	20.00	3	20.00	4	26.67	2	13.33
35 – 65	Average Mastery (AM)	8	53.33	1	6.67	6	40.00	1	6.67
15 – 34	Low Mastery (LM)	4	26.67	0	0.00	4	26.67	0	0.00
Overall Mean		50.00		92.67		53.00		94.33	
Qualitative Description		AM		CAM		AM		CAM	
SD		19.46		12.52		25.06		10.15	
Note. f=frequency; SD=standard deviation; MPS=Mean percentage score									

The traditional approach significantly improved the respondents' spelling proficiency. In the pretest, most learners were at Average Mastery (53.33%) or Low Mastery (26.67%), with none reaching the higher mastery levels. The group obtained an MPS of 50.00, interpreted as Average Mastery. After the intervention, 46.67% of learners reached the Mastered level, and 26.67% reached Closely Approximating Mastery, with no respondents remaining at the Low or Very Low Mastery levels. The posttest MPS increased to 92.67, interpreted as Closely Approximating Mastery, while the lower standard deviation of 12.52 indicated more consistent performance. These findings support Skinner's Behaviorist Theory, which emphasizes repetition, practice, and reinforcement, and Piaget's Cognitive Theory, which connects language development with intellectual growth and conceptual understanding. The structured activities also strengthened phonological awareness, memory, and orthographic knowledge, which are important in spelling acquisition (Reed et al., 2016). This improvement also reflects Krashen's Input Hypothesis, as learners received meaningful, comprehensible tasks in a supportive environment. The AI-based approach also improved spelling proficiency. In the pretest, most learners were classified as Average Mastery (40.00%), Moving Towards Mastery (26.67%), or Low Mastery (26.67%), while only one learner reached Closely Approximating Mastery. The pretest MPS was 53.00, interpreted as Average Mastery. After the intervention, 60.00% of learners reached the Mastered level, and 20.00% reached Closely Approximating Mastery, while only a small number remained at the lower mastery levels. The posttest MPS increased to 94.33, interpreted as Closely Approximating Mastery, and the standard deviation decreased to 10.15, suggesting more consistent performance. These findings support Krashen's Input and Affective Filter Hypotheses because the interactive, adaptive activities provided meaningful input in an engaging, low-stress environment. The results also reflect adaptive learning technologies, which personalize activities and provide immediate feedback based on learner performance (Walgermo et al., 2024). The continuous reinforcement and correction provided through AI-assisted activities are also consistent with Behaviorist Theory.

The comparative results indicate that both traditional and AI-based approaches improved spelling proficiency. However, the AI-based group obtained a slightly higher posttest MPS of 94.33 compared with 92.67 for the traditional group, and a larger proportion reached the Mastered level (60.00%) than in the traditional group (46.67%). The AI-based group also had a lower posttest standard deviation of 10.15 compared with 12.52 in the traditional group, suggesting more consistent learning outcomes. These results suggest an added advantage of AI-based instruction in helping more learners reach higher levels of spelling proficiency. This supports

integrating AI-assisted strategies while maintaining effective traditional teaching practices that emphasize explicit phonological, morphological, and rule-based instruction (Bear et al., 2020; Joshi & Moats, 2021; El Tawab, 2025). Recent studies also support the potential of technology-enhanced spelling instruction. Digital game-based and app-based interventions have been reported to improve spelling accuracy and pattern recall, particularly when they include immediate feedback, adaptive difficulty, and multisensory engagement (Hirsh-Pasek et al., 2020; Mills et al., 2022; Zucker & Light, 2023; Witzel, 2024; Lara, 2025). These findings suggest that AI-assisted and game-based tools may support faster mastery by providing personalized practice and reinforcing correct spelling in an engaging format. However, traditional approaches remain effective, and some studies show that structured rule-based, word study, and systematic phonics-oriented strategies can produce similar or better gains than less guided technology-based approaches when supported by explicit instruction and teacher guidance (Bear et al., 2020; Joshi & Moats, 2021; Hsu et al., 2021; El Tawab, 2025; Saab, 2025).

Contrary findings also show that more advanced spelling methods do not always produce better results than structured traditional practice, particularly for irregular words or when learner-specific factors such as prior literacy level and motivation are not addressed (Hsu et al., 2021; Roscoe & McNamara, 2013). For example, El Tawab (2025) reported that a word study strategy significantly improved Grade 7 learners' spelling proficiency, while the AI-based group outperformed the traditional group using traditional spelling tasks. These findings suggest that the effectiveness of AI-based instruction depends on how it is integrated into the broader spelling curriculum, the quality of teacher guidance, and the AI tool's features, including feedback, task design, and mastery monitoring. Overall, the findings support the value of both approaches in improving spelling proficiency. The traditional approach emphasizes repetition, guided practice, and reinforcement, while the AI-based approach provides adaptive learning, immediate feedback, and learner-centered practice. The results also align with Krashen's theories and Cognitive Theory, which emphasize meaningful input and appropriate learning experiences in language development. Although the AI-based group showed slightly higher and more consistent posttest results, the findings do not establish that AI is superior to traditional instruction. Instead, they suggest that combining effective teacher-led practices with well-designed AI-supported tools may provide a useful approach to improving spelling proficiency and literacy outcomes among learners.

Difference in Vocabulary and Spelling Proficiency Between Pretest and Posttest

Traditional Approach. Table 3 summarizes the mean differences, standard deviations, and statistical significance (e.g., t-test results) in vocabulary and spelling proficiency scores from pretest to posttest, reflecting learning progress under the traditional approach.

Table 3. Pretest–Posttest Differences in Vocabulary and Spelling Proficiency Using the Traditional Approach

Variables	Pretest		Posttest		<i>t</i>	<i>p</i>	Decision	Interpretation
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>				
Vocabulary	29.67	9.72	92.33	16.99	-12.79	< .001	Reject H_0	Significant; Cohen's <i>d</i> = 3.87
Spelling	50.00	19.46	92.67	12.52	-10.04	< .001	Reject H_0	Significant; Cohen's <i>d</i> = 2.59

Note. *M*=mean; *SD*=standard deviation

Table 3 presents the difference in respondents' vocabulary and spelling proficiency before and after implementing the traditional approach. The results show significant improvements in both areas, as indicated by the increase in mean scores and the computed t-values with corresponding p-values at the .001 level of significance. Thus, the null hypothesis was rejected, indicating that the traditional approach significantly improved the respondents' vocabulary and spelling proficiency. For vocabulary, the mean increased from 29.67 (Low Mastery) with a standard deviation of 9.72 in the pretest to 92.33 (Closely Approximating Mastery) in the posttest, although the higher posttest standard deviation of 16.99 indicates greater variation in learners' performance. The computed t-value of -12.79 and p-value of .001 confirm that the improvement was statistically significant. Similarly, spelling proficiency increased from a pretest mean of 50.00 (Average Mastery) with a standard deviation of 19.46 to a posttest mean of 92.67 (Closely Approximating Mastery) with a standard

deviation of 12.52. The computed t-value of -10.04 and p-value of .001 also confirm a statistically significant improvement. These findings support studies emphasizing the effectiveness of traditional teacher-guided instruction in language development. Bear et al. (2020), Joshi and Moats (2021), and Hsu et al. (2021) highlighted the value of explicit instruction, guided practice, corrective feedback, and cumulative review in improving vocabulary and spelling proficiency. The significant gains in the present study demonstrate the value of structured word study, spelling drills, repetition, and reinforcement in helping learners develop vocabulary meanings and spelling patterns. These practices are consistent with Behaviorist Theory, which emphasizes repetition, reinforcement, and habit formation. The findings also support Piaget's Cognitive Theory because structured and sequential activities can strengthen phonological awareness, memory retention, and orthographic knowledge. Likewise, Krashen's Input Hypothesis is reflected in the meaningful and comprehensible language input provided through guided classroom activities and supportive learning experiences.

Although recent studies have highlighted the advantages of AI-based instruction, the present findings show that traditional teaching can also produce high-level learning outcomes. Torres and Kahveci (2025), Liu et al. (2025), and Zhang et al. (2022–2025) reported that AI-supported learning can improve language proficiency through adaptive learning, personalized exercises, and immediate feedback. However, the current findings suggest that well-structured traditional instruction can achieve strong results when learners receive clear modeling, repeated practice, corrective feedback, and teacher guidance. Bear et al. (2020) and Joshi and Moats (2021) further support this, noting that explicit, rule-based instruction may match or exceed the gains of technology-supported approaches when implemented effectively. Overall, the significant improvements in vocabulary and spelling proficiency show that traditional instructional strategies remain reliable and effective, especially in contexts where access to digital technologies may be limited. The findings suggest that teacher-guided activities, repetition, corrective feedback, and cumulative review should remain central to literacy instruction and intervention programs such as ARAL and SARP. The results also have implications for students, teachers, curriculum developers, policymakers, school administrators, and future researchers by highlighting the importance of structured literacy activities and effective teacher guidance. Rather than treating AI and traditional methods as opposing approaches, educators may view them as complementary strategies that can be integrated to meet diverse learner needs. Overall, the findings reinforce that meaningful input, guided reinforcement, and appropriate learning experiences support language learning.

AI-based Approach. Table 4 shows the mean differences, standard deviations, and statistical significance (e.g., paired t-test results) in vocabulary and spelling proficiency from pretest to posttest for respondents using the AI-based approach.

Table 4. Pretest–Posttest Differences in Vocabulary and Spelling Proficiency Using the AI-based Approach

Variables	Pretest		Posttest		<i>t</i>	<i>p</i>	Decision	Interpretation
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>				
Vocabulary	40.67	12.79	93.33	9.57	-14.99	< .001	Reject H_0	Significant; Cohen's $d = 3.87$
Spelling	53.00	25.06	94.33	10.15	-7.25	< .001	Reject H_0	Significant; Cohen's $d = 1.87$

Note. M=mean; SD=standard deviation

The results reveal a highly significant improvement in both vocabulary and spelling proficiency of the AI-based group from pretest to posttest. Vocabulary scores increased from a mean of 40.67 (Average Mastery) to 93.33 (Closely Approximating Mastery), while spelling scores improved from 53.00 (Average Mastery) to 94.33 (Closely Approximating Mastery). The computed t-values (-14.99 for vocabulary and -7.25 for spelling), with p-values less than .001, indicate that these gains are statistically significant, leading to the rejection of the null hypothesis. These findings suggest that the AI-based instructional approach was highly effective in enhancing learners' vocabulary and spelling skills, enabling a substantial shift toward mastery-level performance.

The significant gains observed in the AI-based group align with recent studies showing that AI-enhanced language instruction can produce large, statistically significant improvements in vocabulary and general

language proficiency (Wei, 2023; Torres, 2025; El Tawab, 2025). Meta-analytic and experimental work further indicates that AI-mediated instruction yields substantial overall effects on English learning achievement, including vocabulary, grammar, reading, and writing, outperforming traditional methods across diverse EFL contexts (Jantakoon et al., 2025; Torres, 2025). Classroom-based innovations such as AI-integrated song-based “Sing & Spell with AI” likewise report clear improvements in vocabulary recall and spelling accuracy, particularly among lower-proficiency learners, reinforcing the potential of personalized, feedback-rich AI tools to accelerate mastery in core language subskills (El Tawab, 2025; Singh & Spell with AI, 2025). In contrast, some studies report more modest or mixed advantages, noting that the effectiveness of AI depends heavily on design quality, teacher mediation, and task alignment, and that well-structured traditional methods can still yield strong gains when explicitly scaffolded (Bear et al., 2020; Joshi & Moats, 2021; Hsu et al., 2021). These contrasting findings suggest that while AI-based instruction can strongly boost vocabulary and spelling proficiency, it works best when integrated with strong traditional teaching practices to support sustained, high-level mastery.

Pre-test Difference in Vocabulary and Spelling Proficiency

Table 5 presents the mean scores, standard deviations, and inferential statistics (e.g., t-test results) comparing vocabulary and spelling proficiency between respondents in the traditional and AI-based approaches at pretest.

Table 5. Pre-test Difference in Vocabulary and Spelling Proficiency Using the Traditional and AI-based Approach

Variables	Traditional Approach		AI-based Approach		<i>t</i>	<i>p</i>	Decision	Interpretation
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>				
Vocabulary	29.67	9.72	40.67	12.80	-2.65	.013	Reject H_0	Significant; Cohen's $d = 0.97$
Spelling	50.00	19.46	53.00	25.06	-.37	.717	Do not Reject H_0	Not Significant; Cohen's $d = 0.13$

Note. M=mean; SD=standard deviation

The pretest results established the learners’ baseline proficiency, prior knowledge, readiness, and initial differences before the intervention. In the traditional approach, respondents obtained a mean score of 29.67 ($SD = 9.72$) in vocabulary and 50.00 ($SD = 19.46$) in spelling, indicating lower vocabulary proficiency but relatively comparable spelling proficiency. The lower vocabulary score may suggest limited prior exposure to vocabulary-enrichment activities. These findings are supported by Reed et al. (2016), who explained that vocabulary and spelling development depend on phonological awareness, orthographic knowledge, memory, and repeated exposure to words. Since vocabulary and spelling are commonly included in regular classroom instruction, learners may have developed similar foundational literacy skills before the study. In the AI-based approach, respondents obtained a mean score of 40.67 ($SD = 12.80$) in vocabulary and 53.00 ($SD = 25.06$) in spelling. These results show that the AI-based group had higher vocabulary and slightly higher spelling proficiency than the traditional group before the intervention. The significant vocabulary difference ($p = .013$) indicates that the AI-based group already had stronger vocabulary knowledge at baseline. This finding may be explained by Krashen’s Input Hypothesis, which emphasizes meaningful and comprehensible input, and Piaget’s Cognitive Theory, which highlights active cognitive engagement and meaningful learning experiences. AI-supported environments may provide interactive activities, personalized tasks, immediate feedback, and repeated exposure that support vocabulary learning. The slightly higher spelling score may also reflect adaptive learning and immediate feedback, which support repeated practice, orthographic awareness, and memory reinforcement (Reed et al., 2016).

The pretest findings are also consistent with studies on AI-enhanced language learning. Torres (2025) reported that AI-powered learning tools can improve vocabulary acquisition, retention, and motivation through adaptive exercises and immediate feedback, while Liu et al. (2025) found that AI-enhanced systems support language proficiency through individualized and interactive learning environments. Mugableh (2024) also reported gains in vocabulary knowledge through personalized and engaging AI-generated activities. Similarly, Rahman (2025), Lara (2025), and Witzel (2024) found that AI-assisted spelling interventions can improve spelling accuracy and

learner engagement through interactive and scaffolded activities. These studies support the potential role of AI-based instruction in language learning through personalization, immediate feedback, engagement, and adaptive support. Table 5 presents the comparative pretest results of the traditional and AI-based groups. Vocabulary proficiency differed significantly, with the AI-based group achieving a higher mean score of 40.67 compared with 29.67 for the traditional group. The computed t -value of -2.65 and p -value of .013 indicate that the groups were not equivalent in vocabulary proficiency before the intervention. In contrast, no significant difference was found in spelling proficiency, as shown by the close mean scores of 50.00 for the traditional group and 53.00 for the AI-based group, with a p -value of .717. This indicates that the groups had comparable spelling proficiency at baseline.

Consider the baseline vocabulary difference when interpreting the posttest results, because learners with higher initial proficiency may show different levels of improvement. Recent studies emphasize the importance of baseline equivalence and recommend statistical control or matching procedures when initial language proficiency differs between groups (Wei, 2023; Jantakoon et al., 2025; El Tawab, 2025). Similarly, Torres (2025) and Singh & Spell with AI (2025) noted that learners with higher initial vocabulary or spelling proficiency may show larger gains. The findings also support Piaget's Cognitive Theory, as previous experiences and cognitive engagement can influence knowledge development, and Krashen's Input Hypothesis, which emphasizes the role of meaningful and comprehensible input. Overall, the pretest results show that the AI-based group had significantly higher vocabulary proficiency, while both groups had comparable spelling proficiency before the intervention. These results highlight the importance of pretest assessment in establishing baseline conditions and accounting for initial differences when evaluating an intervention's effects.

Post-test Difference in Vocabulary and Spelling Proficiency

Table 6 presents the mean scores, standard deviations, and inferential statistics (e.g., t -test results) comparing vocabulary and spelling proficiency of the respondents between the traditional approach and AI-based approach at posttest. The posttest results emphasized learners' performance after the intervention, highlighting the effectiveness of the traditional and AI-based approaches in improving vocabulary and spelling proficiency. The findings provide a basis for evaluating learning gains, determining the impact of the interventions, and informing recommendations for language instruction and literacy development.

Table 6. Post-test Difference in Vocabulary and Spelling Proficiency Using the Traditional and AI-based Approach

Variables	Traditional Approach		AI-based Approach		t	p	Decision	Interpretation
	M	SD	M	SD				
Vocabulary	92.33	16.99	93.33	9.57	-0.20	.844	Do not Reject H_0	Not Significant; Cohen's $d = 0.07$
Spelling	92.67	12.52	94.33	10.15	-0.40	.692	Do not Reject H_0	Not Significant; Cohen's $d = 0.15$

Note. M =mean; SD =standard deviation

The posttest results show that the traditional approach led to high levels of vocabulary and spelling proficiency. Respondents earned a mean score of 92.33 ($SD = 16.99$) in vocabulary and 92.67 ($SD = 12.52$) in spelling. These high scores suggest that structured instruction, guided practice, repetition, and reinforcement activities helped improve learners' language skills. The findings support Skinner's Behaviorist Theory, which explains that language learning develops through repeated exposure, practice, imitation, and reinforcement. Traditional activities such as drills, spelling exercises, vocabulary discussions, and teacher-guided instruction may have strengthened learners' retention and mastery of words. The findings also align with foundational literacy principles and studies by Bear et al. (2020), Joshi and Moats (2021), and Hsu et al. (2021), which emphasized the value of explicit, systematic, guided, and rule-based instruction in improving vocabulary and spelling proficiency. The AI-based group also achieved high posttest proficiency, with a mean score of 93.33 ($SD = 9.57$) in vocabulary and 94.33 ($SD = 10.15$) in spelling. These results suggest that AI-supported activities enhanced learners' vocabulary and spelling through interactive, adaptive, and feedback-rich learning experiences. The findings support Krashen's Input Hypothesis, which emphasizes meaningful and comprehensible input in

engaging learning environments, as well as Piaget's Cognitive Theory, which highlights active engagement and meaningful interaction with learning materials. AI-based activities provided immediate feedback, repeated exposure, personalized tasks, and opportunities for self-paced practice. The results are also consistent with Torres (2025) and Liu et al. (2025), who reported benefits of AI-powered tools in vocabulary and language learning, and with Rahman (2025), Lara (2025), and Witzel (2024), who found improvements in spelling accuracy and learner engagement through AI-assisted activities.

The comparative posttest results presented no significant difference between the traditional and AI-based groups in either vocabulary or spelling proficiency. For vocabulary, the traditional group obtained a mean of 92.33, while the AI-based group obtained 93.33, with $p = .844$. For spelling, the means were 92.67 and 94.33, respectively, with $p = .692$. Although the AI-based group scored slightly higher, the differences were not statistically significant. Both groups therefore achieved comparable high levels of proficiency, reaching Closely Approximating Mastery (CAM). These findings suggest that both approaches can effectively support vocabulary and spelling development and that either approach, or a combination of both, may be used in instructional practice. The absence of a significant posttest difference is consistent with findings that AI may be more useful as a complementary tool than as a replacement for traditional instruction. Ya'u and Mohammed (2025) found that AI tools can enhance academic literacy, particularly when integrated with traditional teaching rather than used as a replacement. Similarly, "Artificial Intelligence in the Teaching of English Grammar" (2025) reported learning gains from AI-based instruction while noting teacher resistance and differences in implementation, suggesting that contextual factors may influence outcomes. However, other studies have reported greater gains from AI-based instruction. Mugableh (2024) found improvements in vocabulary size, lexical depth, and spelling accuracy through ChatGPT-generated exercises, while El Tawab (2025) and Saab (2025) reported positive effects of AI-based vocabulary instruction among EFL learners. These findings suggest that AI's effectiveness may depend on factors such as grade level, instructional design, intervention duration, and the specific learning tasks used. Overall, the posttest results indicate that both traditional and AI-based approaches successfully improved learners' vocabulary and spelling proficiency. The AI-based group showed slightly higher mean scores in both areas, but these differences were not statistically significant, indicating comparable posttest performance. The findings therefore do not establish that AI-based instruction is superior to traditional teaching. Instead, they support the view that AI can complement effective teacher-guided instruction by providing interactive activities, immediate feedback, and additional practice opportunities. The results further suggest that combining appropriate AI-supported tools with established traditional teaching practices may be a practical approach to strengthening language learning and literacy development.

Analysis of Covariance (ANCOVA) of Posttest Vocabulary and Spelling Scores

Table 7 presents the ANCOVA results using the corresponding pretest score as the covariate. For vocabulary, the adjusted posttest means were 93.33 for the traditional group and 92.34 for the AI-based group. The adjusted between-group effect was not significant, $F(1,27) = 0.03$, $p = .863$, partial $\eta^2 = .001$. For spelling, the adjusted posttest means were approximately 93.05 for the traditional group and 93.95 for the AI-based group. The adjusted between-group effect was also not significant, $F(1,27) = 0.06$, $p = .807$, partial $\eta^2 = .002$. These findings indicate that, after controlling for initial proficiency, neither instructional approach demonstrated a statistically significant advantage over the other.

Table 7. Analysis of Covariance (ANCOVA) of Posttest Vocabulary and Spelling Scores Controlling for Pretest Scores

Outcome	Adjusted Traditional M	Adjusted AI-based M	F(1,27)	p	Partial η^2	Cohen's d	Interpretation
Vocabulary	93.33	92.34	0.03	.863	.001	0.07	Not significant
Spelling	93.05	93.95	0.06	.807	.002	0.15	Not significant

CONCLUSION

Based on the findings, both traditional and AI-based instructional approaches significantly improved Grade 9 learners' vocabulary and spelling proficiency from pretest to posttest. The AI-based group had higher baseline vocabulary scores, indicating that the groups were not fully equivalent before the intervention. After controlling for baseline proficiency through ANCOVA, however, no statistically significant between-group difference was found in vocabulary, $F(1,27) = 0.03$, $p = .863$, or spelling, $F(1,27) = 0.06$, $p = .807$. The very small ANCOVA effect sizes further indicate that the adjusted differences between instructional approaches were negligible. Therefore, the results do not support a claim that AI-based instruction was superior to traditional instruction in this study. Instead, both approaches appear capable of producing substantial short-term gains, and AI may be considered a supplementary tool within teacher-guided instruction.

Limitations and Recommendations

The study was limited to 30 Grade 9 learners from intact sections in one secondary school and used a 10-day intervention with an immediate posttest. Thus, the findings should be interpreted as evidence of short-term learning outcomes within the study setting rather than broad evidence of generalizability or long-term retention. The study also did not include a formal qualitative component such as student or teacher interviews.

Future studies should include learners from multiple schools and sections to strengthen generalizability, and should administer a delayed posttest approximately 4–6 weeks after the intervention to examine retention. Student and teacher interviews or focus-group discussions may also be incorporated to explore engagement, usability, perceived benefits, and implementation challenges. Where baseline differences exist, ANCOVA or another appropriate covariate-adjusted analysis should be used with individual-level scores, subject to the relevant statistical assumptions.

REFERENCES

1. Bear, D. R., Invernizzi, M., Templeton, S., & Johnston, F. (2020). Words their way: Word study for phonics, vocabulary, and spelling instruction (7th ed.). Pearson. <https://tinyurl.com/yyepx46y>
2. Department of Education. (2020). Most essential learning competencies (MELCs). <https://lrmds.deped.gov.ph/detail/18275>
3. El Tawab, D. S. M. I. (2025). The effect of artificial intelligence (AI) approach to develop EFL vocabulary among primary pupils. *Journal of Educational Research in Middle East and Africa*, 1, 112. <https://doi.org/10.21608/ssj.2025.443758>
4. Jantakoon, T., Jantakun, T., Jantakun, K., Pongpanich, W., Pasmala, R., Wannapiroon, P., & Nilsook, P. (2025). The effectiveness of artificial intelligence in English instruction for speaking and listening skills: A meta-analysis. *Contemporary Educational Technology*, 17(4), Article ep596. <https://doi.org/10.30935/cedtech/17310>
5. Joshi, R. M., & Moats, L. C. (2021). Reading, writing, and spelling for students with learning disabilities: A clinician's guide. Brookes Publishing.
6. Liu, J., Sihes, A. J. B., & Lu, Y. (2025). How do generative AI tools and large language models (LLMs) influence language learners' critical thinking in EFL education? A systematic review. *Smart Learning Environments*, 12, Article 48. <https://doi.org/10.1186/s40561-025-00406-0>
7. Mekheimer, K. (2025). The role of AI-powered applications in enhancing language skills: Global and local perspectives. *International Journal of Educational Innovation*, 18(2), 44–59.
8. Mugableh, A. I. (2024). The impact of ChatGPT on the development of vocabulary knowledge of Saudi EFL students. *Arab World English Journal*, 1(1), 265–281. <https://doi.org/10.24093/awej/ChatGPT.18>
9. Reed, D. K., Petscher, Y., & Foorman, B. R. (2016). The contribution of vocabulary knowledge and spelling to the reading comprehension of adolescents who are and are not English language learners. *Reading and Writing*, 29(4), 633–657. <https://doi.org/10.1007/s11145-015-9619-3>
10. Rosqueta, K. (2025). AI in Philippine classrooms: Opportunities and barriers. *Philippine Journal of Educational Research*, 5(1), 55–73.

11. Saab, J. (2025). The impact of artificial intelligence tools on vocabulary learning of EFL university students. *Al-Baath Journal for Arts, Humanities and Social Sciences*, 47, 173–192. <https://asjp.cerist.dz/en/article/273164>
12. Singh & Spell with AI. (2025). Sing & spell with AI: Enhancing vocabulary acquisition and spelling accuracy in Year 2 pupils. *International Journal of Research and Innovation in Social Science*, 9(2), 215–218.
13. Stevens, J. P. (2025). Artificial intelligence in the teaching of English grammar. *International Journal of Research and Innovation in Applied Science*, X(VI), 878–896. <https://doi.org/10.51584/IJRIAS.2025.10060066>
14. Torres, P. J., & Kahveci, Y. E. (2025). Effectiveness of artificial intelligence (AI) in language teaching. *Computers & Education: Artificial Intelligence*, 9, Article 100522. <https://doi.org/10.1016/j.caeai.2025.100522>
15. Walgermo, B. R., Foldnes, N., Uppstad, P. H., Bakken, A. M., & Lundetræ, K. (2024). Development and deployment of an adaptive national elementary reading screening test. *Frontiers in Education*, 9, Article 1331777. <https://doi.org/10.3389/feduc.2024.1331777>
16. Wei, L. (2023). Artificial intelligence in language instruction: Impact on English learning achievement, L2 motivation, and self-regulated learning. *Frontiers in Psychology*, 14, Article 1261955. <https://doi.org/10.3389/fpsyg.2023.1261955>
17. Ya'u, M. S., & Mohammed, M. S. (2025). AI-assisted writing and academic literacy: Investigating the dual impact of language models on writing proficiency and ethical concerns in Nigerian higher education. *International Journal of Education and Literacy Studies*, 13(2), 593–604. <https://doi.org/10.7575/aiac.ijels.v.13n.2p.593>