Exploring the Use of AI-Generated Contextual Sentences for English Vocabulary Acquisition Among Malaysian ESL Learners

Lavaneesh Rama, Samuel Manickam, Yuvavatania Davudars, Mohd Haniff Mohd Tahir

Faculty of Languages and Communication, University Pendidikan Sultan Idris, Tg. Malim, Perak, Malaysia

DOI: https://dx.doi.org/10.47772/IJRISS.2026.10100119

Received: 04 January 2026; Accepted: 09 January 2026; Published: 24 January 2026

ABSTRACT

This study investigates the efficacy of AI-generated contextual sentences in enhancing vocabulary acquisition among Malaysian ESL learners. Vocabulary mastery is pivotal for language proficiency, yet rote learning methods often fail to provide the necessary context for long-term retention. Adopting a mixed-method research design, this study utilises methodological triangulation comprising pre- and post-tests, semi-structured teacher interviews, and student interviews. The study involved 10 Form Four ESL students and 3 English teachers from a Malaysian secondary school in Ipoh, Perak, and evaluated the effectiveness of AI-customised vocabulary exercises tailored to learner interests. Quantitative data were analysed using pre- and post-test comparisons, which revealed a clear improvement in vocabulary definition accuracy and contextual usage following the intervention. Qualitative findings further indicated heightened learner engagement and motivation. However, discrepancies between written vocabulary mastery and oral fluency were observed. These findings provide important implications for the pedagogical integration of Generative AI in English Language Teaching (ELT).

Keywords: AI-generated sentences, vocabulary acquisition, ESL learners, contextual learning.

INTRODUCTION

Vocabulary acquisition is a cornerstone of linguistic proficiency. In the Malaysian education context, despite over a decade of formal English instruction, many secondary school students remain "lexically compromised," often not progressing at intermediate levels of proficiency. This stagnation is frequently attributed to the pedagogical limitations of the classroom environment, where vocabulary is treated as a static list of items to be memorised rather than a dynamic system of meaning to be acquired. Also, Malaysian ESL learners frequently struggle to retain and apply new lexis, often due to decontextualised pedagogical approaches such as rote memorisation and direct translation (Nation, 2022).

Current Second Language Acquisition (SLA) research suggests that vocabulary is best acquired when presented in meaningful, authentic contexts. Artificial Intelligence (AI), specifically Large Language Models (LLMs) like ChatGPT, offers a novel solution by generating personalised, context-rich sentences based on learner interests. This study addresses the gap in empirical research regarding AI-driven vocabulary learning by exploring three key areas: the impact on retention, teacher perceptions of pedagogical effectiveness, and student engagement levels.

Research Objectives

The present study, entitled *"Exploring the Use of AI-Generated Contextual Sentences for English Vocabulary Acquisition among Malaysian ESL Learners"*, aims to address this gap by pursuing the following objectives:

1. To examine the effectiveness of AI-generated contextual sentences in improving vocabulary retention and contextual usage among Form 4 ESL learners.
2. To explore teachers' perceptions of pedagogical effectiveness and classroom dynamics associated with the use of AI-generated contextual sentences.

3. To investigate students' perceptions of their engagement, motivation, and learning relevance when using AI-generated vocabulary exercises.

Research Questions (RQ)

In line with the above objectives, this study seeks to answer the following research questions:

1. Does the use of AI-generated contextual sentences lead to a significant improvement in vocabulary retention and usage among Form 4 ESL learners?

2. What are teachers' perceptions regarding the pedagogical effectiveness and classroom dynamics when utilising AI-generated contextual sentences?

3. How do students perceive their engagement, motivation, and learning relevance when interacting with AI-generated vocabulary exercises?

By addressing these questions, this research aims to contribute to the limited body of empirical work on AI-driven vocabulary learning and acquisition.

LITERATURE REVIEW

Researchers have long sought to understand how people acquire new vocabulary in a second language. Traditional methodologies have predominantly relied on "intentional learning" strategies, specifically, rote memorisation of word lists, bilingual translation, and decontextualised drilling. While these methods facilitate rapid initial mapping of form to meaning, they often result in "inert knowledge" vocabulary that learners can recognize in isolation but fail to retrieve or apply in communicative contexts (Nation, 2022).

The "Context Gap" in Vocabulary Pedagogy

Recent studies indicate that for vocabulary to move from receptive to productive knowledge, it must be encountered in "rich" contexts that provide semantic, syntactic, and pragmatic clues (Schmitt, 2023). The absence of such context in standard textbooks creates a "context gap," where learners understand the definition of a word like meticulous but cannot distinguish its usage from synonyms like careful or detailed. This gap is particularly pronounced in the post-pandemic era, where digital literacy has outpaced traditional pedagogical evolution (Konyrova, 2024).

In parallel, technology has transformed language learning. AI tools, including chatbots and adaptive platforms, are increasingly integrated into classrooms. Unlike traditional methods, AI-powered tools can offer vocabularies embedded in dynamic, personalised contexts generating sentences based on learner interests or simulating real-life dialogues. Despite these developments, few studies have specifically investigated the effectiveness of AI-generated contextual sentences in supporting vocabulary acquisition, with most research focusing on general digital tools.

Generative AI as a "Digital Scaffold" (Vygotsky's ZPD)

This study is theoretically anchored in Vygotsky's Sociocultural Theory, specifically the Zone of Proximal Development (ZPD). Vygotsky posited that cognitive development is socially situated and mediated by tools and language. In the modern context, Generative AI (GenAI) assumes the role of a "Digital More Knowledgeable Other" (MKO). Unlike static dictionaries, GenAI can generate sentences that are slightly above the learner's current proficiency but rooted in familiar contexts, effectively establishing a personalised ZPD (Hamedi et al., 2025).

Recent scholarship by Munshi (2025) and Wang et al. (2024) suggests that AI tools function as "adaptive scaffolds." By tailoring the complexity of the input to the learner's level, AI provides the necessary "i+1" (comprehensible input) that prevents both boredom and cognitive overload. This challenges the traditional CALL (Computer-Assisted Language Learning) paradigm, which relied on pre-programmed, static responses. GenAI's ability to generate infinite variations of context facilitates "receptive repetition" the requirement that a learner

must encounter a word 5–16 times in different contexts to fully acquire it without the fatigue associated with repetitive drills (Zhao, 2023).

The Personalisation-Affect Connection

Beyond cognitive scaffolding, AI addresses the affective dimension of learning. The "Affective Filter Hypothesis" suggests that anxiety impedes language acquisition. Personalised learning, where content is tailored to learner interests (e.g., sports, music, gaming), has been shown to lower this filter. Alharbi and Khalil (2023) argue that when AI generates content relevant to a student's "lived reality," it triggers the "self-reference effect" a psychological phenomenon where information linked to the self is encoded more deeply in memory. This represents a significant departure from "one-size-fits-all" textbook scenarios, positioning AI not just as a tutor, but as a personalised content creator that validates the learner's identity (Majid & Matore, 2024).

METHODOLOGY

This study adopts a Mixed Method Research Design, specifically utilising a triangulation approach to ensure the validity and reliability of the findings. By integrating both quantitative and qualitative data, the study provides a holistic view of the research problem. The quantitative aspect measures the objective improvement in vocabulary retention through statistical analysis of test scores, while the qualitative aspect explores the subjective experiences, engagement levels, and pedagogical perceptions of both students and teachers. This convergent design allows for the corroboration of results, where numerical data is enriched and explained by narrative insights.

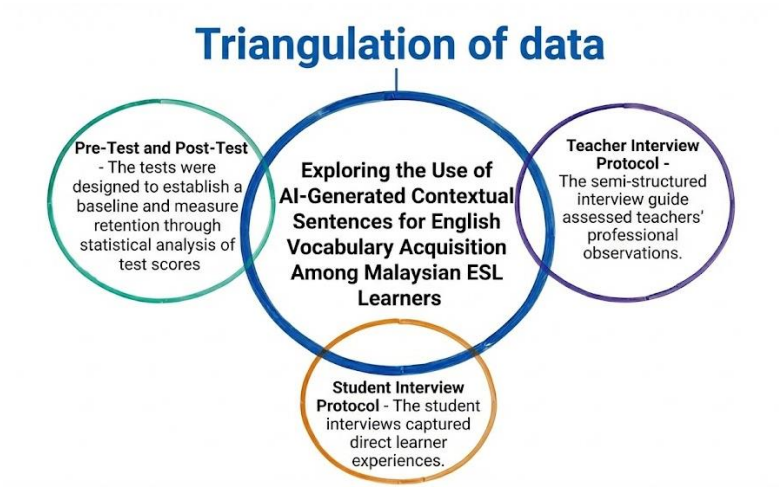

Figure 1: Triangulation of Data Using Three Research Instruments

The study centres on a purposive sample of 10 Form 4 English students enrolled at a secondary school in Ipoh, Perak. These students possess an intermediate level of English proficiency. Additionally, 3 English teachers from the same institution participated in the study to provide expert observational data regarding student performance and classroom dynamics.

To ensure robust triangulation, three distinct instruments were utilised. Below are the specific operational questions/items for each instrument:

Pre-Test and Post-Test

The tests were designed to establish a baseline and measure retention. Three specific item types used in the test instrument were:

- Item 1 (Multiple Choice - Definition): "Which of the following is the best definition of the word 'meticulous'? (a) Careless, (b) Detailed, (c) Lazy, (d) Quick."
- Item 2 (Word Usage in Context): "The scientist's _____ led to groundbreaking discoveries in the field of genetics. (Options: negligence, curiosity, reluctance, silence)."

- Item 3 (Sentence Generation): "Write a complete sentence using the word 'euphoria' to demonstrate your understanding of its meaning."

Teacher Interview Protocol

The semi-structured interview guide assessed teachers' professional observations. Three key questions asked during the interview were:

- **Question 1 (General Impressions):** "What are your general thoughts on the use of AI-generated contextual sentences for teaching vocabulary compared to traditional methods?"
- **Question 2 (Student Motivation):** "Have you noticed a specific change in the motivation of your students to learn new vocabulary since the introduction of AI-generated sentences?"
- **Question 3 (Classroom Performance):** "Have you observed any tangible improvements in the way students use the new vocabulary in their written work or class discussions?"\

Student Interview Protocol

The student interviews captured direct learner experiences. Three key questions asked during the interview were:

- **Question 1 (Engagement):** "Did the personalised AI-generated sentences make you feel more motivated to learn vocabulary compared to standard textbook lists?"
- **Question 2 (Relevance):** "Were the AI-generated sentences relevant to your personal interests or daily life, and did this help you understand the words better?"
- **Question 3 (Retention):** "Do you feel you are able to remember the vocabulary words better because they were presented in these specific sentences?"

The interview transcripts were analysed using thematic analysis (Braun & Clarke, 2006).

The study followed a sequential procedure. First, the Pre-Test was administered to all participants to gather baseline quantitative data. Following this, the intervention phase was implemented, where students utilised AI-generated contextual sentences tailored to their interests for vocabulary learning. Upon completion of the intervention, the Post-Test was administered. Finally, in-depth interviews were conducted with both teachers and students to gather qualitative feedback. Quantitative data from the pre- and post-tests were analysed to determine improvement in scores. Qualitative data from interviews were transcribed and thematically analysed to identify recurring patterns regarding engagement, motivation, and challenges.

Research Design and Intervention Timeline

This research adopted a mixed-method approach, aiming to comprehensively evaluate the effectiveness of AI-generated contextual sentences for Malaysian ESL learners. The intervention unfolded over a structured four-week period, methodically guiding students from basic vocabulary recognition towards the practical application of new words within meaningful contexts. In the first week, all 10 Form 4 participants underwent a pre-test designed to establish a clear baseline for their vocabulary knowledge and their ability to use newly introduced words in context. This initial assessment was crucial for accurately measuring progress and identifying specific areas of need for each learner.

During the second and third weeks, students engaged in a series of vocabulary exercises generated by a sophisticated Large Language Model (LLM). These exercises were thoughtfully tailored to align with the students' personal interests such as sports, music, and other relevant themes to ensure that the learning material was not only educational but also engaging and relatable. This personalised approach aimed to enhance motivation and retention, making vocabulary acquisition more meaningful, as students could see the direct relevance to their everyday experiences.

In the final week, the intervention concluded with a comprehensive post-test and in-depth semi-structured interviews. Both students and teachers participated in these interviews, providing rich, qualitative feedback that complemented the quantitative test results. This dual-method approach allowed for a more nuanced understanding of the intervention's impact, capturing both measurable gains in vocabulary and subtler insights into learner engagement, challenges faced, and perceptions of the AI-generated materials. Collecting data from multiple stakeholders also ensured that the findings accurately reflected the realities of classroom implementation, rather than just theoretical outcomes.

AI Prompt Engineering and Teacher Mediation

Teachers played an indispensable role as "pedagogical mediators," ensuring that the AI's output was not only accurate but also developmentally appropriate and culturally relevant for the students. Students did not interact directly with the AI; instead, teachers acted as intermediaries, carefully crafting prompts to generate content that was strictly aligned with the students' current proficiency and learning needs. For example, a teacher might input: "Generate three simple contextual sentences for the word 'meticulous' suitable for a Malaysian Form 4 ESL student at a B1 CEFR level, focusing on a sports-related scenario." By doing so, teachers were able to closely control the complexity and relevance of the material, maintaining it within the students' Zone of Proximal Development (ZPD) and avoiding the pitfalls of overwhelming or inappropriate content.

This approach effectively bridged the "context gap" frequently observed in conventional ESL textbooks, which often lack sufficient real-world relevance or adaptability to individual learners' backgrounds. The AI, in this process, served as a "Digital More Knowledgeable Other" (MKO), offering flexible and immediate input that was responsive to the teacher's pedagogical intentions. The collaboration between teacher and AI ensured that students received support that was both scaffolded and contextually grounded, promoting deeper understanding and more confident use of new vocabulary.

Ethical Considerations and Sample Limitations

Strict ethical protocols were maintained throughout the study, given the involvement of minors. Informed consent was obtained from parents or guardians for all student participants, ensuring that ethical standards for research with young people were rigorously upheld. To safeguard privacy and confidentiality, all student data was anonymised using codes such as Student 1, Student 2, and so on, preventing any identification of individual participants in the analysis or reporting of findings.

Nonetheless, the research is subject to several notable limitations. The sample was limited to just 10 students and 3 teachers from a single secondary school in Ipoh, Perak, which means the findings reflect a very specific context and may not be readily generalisable to the broader population of Malaysian ESL learners. The small sample size also limits the statistical power of the quantitative findings and the diversity of perspectives captured in the qualitative data. For future research, it will be important to expand the participant pool to include students from a wider range of schools, regions, and backgrounds. This broader approach would help determine whether the promising outcomes observed in this preliminary study can be replicated and sustained on a larger scale, ultimately contributing to a more comprehensive understanding of AI-assisted vocabulary learning in diverse Malaysian ESL contexts.

FINDINGS AND ANALYSIS

Research Question 1: Does the use of AI-generated contextual sentences lead to a significant improvement in vocabulary definitions and usage in context among Form 4 ESL learners?

The quantitative analysis of the pre- and post-test instruments demonstrated a clear progression in students' ability to define, recognise, and apply new vocabulary. The data revealed marked improvements across definition identification and context application.

Table 1 Participants' pre-test and post-test scores and the margin of improvement

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume IX Issue X October 2025

Participant	Pre-Intervention Score (%)	Post-Intervention Score (%)	Margin of Improvement (%)
Student 1	45	90	45
Student 2	52	95	43
Student 3	30	92	62
Student 4	40	88	48
Student 5	55	94	39
Student 6	48	91	43
Student 7	35	93	58
Student 8	60	96	36
Student 9	50	89	39
Student 10	42	90	48

Table 1 provides a breakdown of individual learner performance, highlighting the specific gains made by each participant. The "Margin of Improvement" column reveals that every participant achieved a double-digit increase in their scores, with improvements ranging from 36% to 62%. Notably, Participant 3, who entered with the lowest baseline score of 30%, demonstrated the most significant growth, achieving a 62% increase to reach a post-test score of 92%. This suggests the intervention was particularly effective for students with lower initial proficiency. Even Participant 8, who started with the highest pre-intervention score of 60%, improved by 36%, indicating that the AI-generated context methodology offered value to advanced learners as well as those who were struggling.

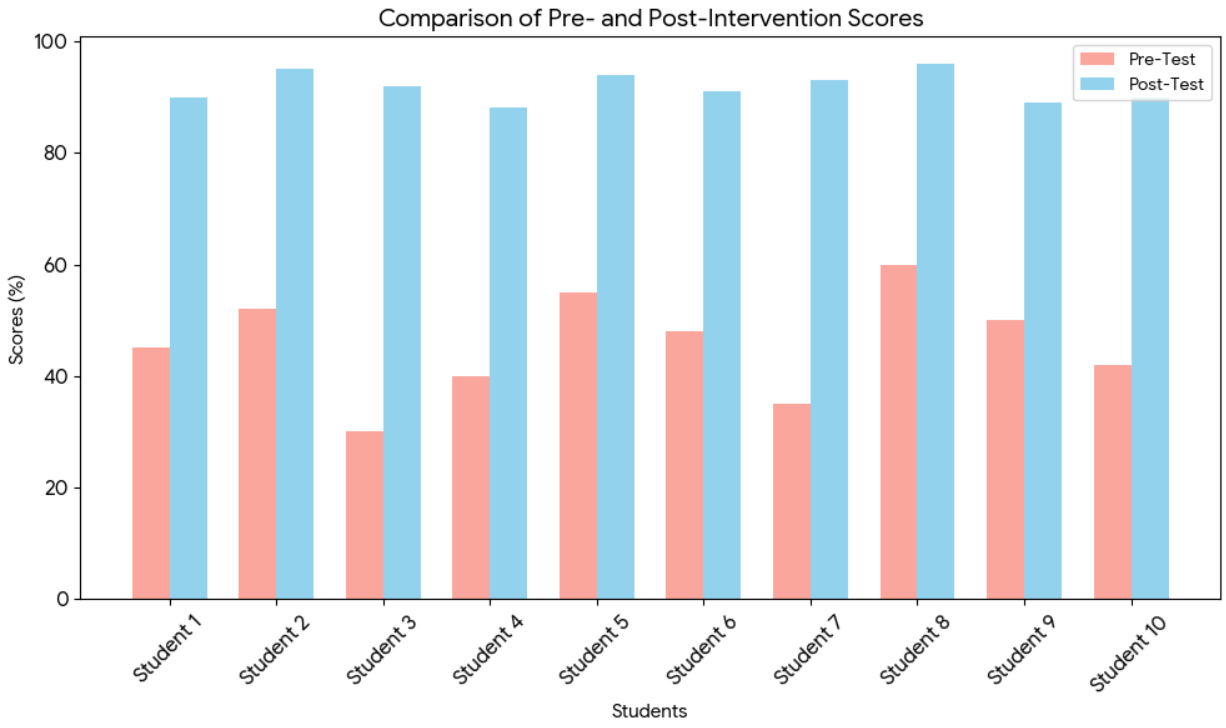

Figure 1: Scores of pre-intervention tests and post-intervention tests

Figure 1 visually illustrates the stark contrast between pre-intervention and post-intervention performance. The histogram displays two distinct, non-overlapping clusters of data. The "Pre-Test" bars (in orange) are concentrated at the lower end of the y-axis, generally fluctuating between the 30% and 60% marks. In contrast, the "Post-Test" bars (in blue) show a uniform upward shift, consistently reaching above the 88% threshold. This visual gap signifies a complete mastery shift; the lowest score in the post-test phase is still considerably higher than the highest score in the pre-test phase. The consistent height of the blue bars suggests that the intervention not only improved scores but also helped normalise performance across the class, bringing all students to a high level of competency regardless of their starting point. The mean score was also calculated to analyse the data gained from the pre- and post-intervention test. The mean was calculated to find the average score of the pupils. The calculation was done by using the formula below.

$$\bar{x} = \frac{\sum x}{n}$$

Application to Data (n= 10)

- Mean of Pre-Intervention Scores:

$$\bar{x}_{pre} = \frac{45 + 52 + 30 + 40 + 55 + 48 + 35 + 60 + 50 + 42}{10} = \frac{457}{10} = \mathbf{45.7\%}$$

- Mean of Post-Intervention Scores:

$$\bar{x}_{post} = \frac{90 + 95 + 92 + 88 + 94 + 91 + 93 + 96 + 89 + 90}{10} = \frac{918}{10} = \mathbf{91.8\%}$$

- Difference of Means (Improvement):

$$\text{Difference} = \bar{x}_{post} - \bar{x}_{pre} = 91.8\% - 45.7\% = \mathbf{46.1\%}$$

Comparing the mean scores clearly shows how much the students improved. The pre-intervention mean of 45.7% indicates that the average participant initially struggled to understand the vocabulary and was performing below the passing mark. However, the post-intervention mean rose significantly to 91.8%, showing that the class as a whole reached a high level of mastery. The difference of 46.1% is a very strong result. It highlights that the average score did not just increase slightly; it effectively doubled. This confirms that the intervention successfully moved the class average from a failing level to an excellent one, proving the method was highly effective.

Improvement in Definitions and Retention

In the pre-test phase, a significant portion of the participants struggled to accurately define the target vocabulary. For instance, data showed that Student 3 incorrectly defined the word 'meticulous' as "Careless," while Student 7 defined it as "Quick." Following the intervention using AI-generated sentences, the post-test results indicated a dramatic shift. 100% of the students, including the previously struggling Student 3 and Student 7, correctly identified the definition as "Detailed." This suggests that the AI-generated contexts helped solidify the semantic meaning of the words in the students' memory.

Enhanced Contextual Usage

Beyond simple definitions, the tests measured the ability to use words in context. In the "Word Usage in Context" section of the pre-test, Student 4 initially selected "negligence" for a sentence that required the word "curiosity." In the post-test, this error was corrected, with the student selecting the appropriate term. This statistical shift indicates that the intervention did not just aid in rote memorisation but improved lexical accuracy and the ability to distinguish between nuanced meanings within a sentence structure.

Research Question 2: What are the teachers' perceptions regarding the pedagogical effectiveness, student motivation, and challenges when utilising AI-generated contextual sentences in the classroom?

Thematic analysis of the teacher interviews revealed three major themes: (1) increased student motivation through personalisation, (2) improved written proficiency versus oral hesitancy, and (3) the need for differentiation.

Increased Student Motivation through Personalisation

Teachers unanimously observed a shift in student attitude toward vocabulary learning. Teacher 1 described students as "excited" to tackle exercises that reflected their personal lives, noting, "The students started incorporating the new vocabulary into their writing; they did not just memorise the words; they were able to use them in a meaningful way." Teacher 2 echoed this, highlighting that the personalisation such as sentences about football for a sports fan transformed the learning from a mechanical chore into a relevant activity. Teacher 3 added that bridging the gap between lists and actual usage empowered students to take responsibility for their learning.

Improved Written Proficiency versus Oral Hesitancy

A key finding from the teacher interviews was the divergence between written and spoken performance. While Teacher 2 identified clear progress in the quality of written expression, nothing words were applied "in ways that made sense within the context of their own writing," oral fluency lagged behind. Teacher 3 observed, "They were able to use the words in written tasks with confidence, but when it came to speaking, they sometimes struggled." This suggests the AI tool successfully built knowledge and written confidence, but additional interventions are needed for oral fluency.

Challenges with Vocabulary Complexity

Despite the success, teachers noted that the "one-size-fits-all" approach of the AI generation had limitations. Teacher 2 stated, "Some of the words were a bit advanced for students at the lower levels." This highlights a need for differentiation within the AI prompts to ensure inclusivity for students with varying proficiency levels.

Research Question 3: How do students perceive their own engagement, motivation, and the relevance of the content when interacting with AI-generated vocabulary exercises?

Thematic analysis of the student interviews focused on their direct experience, revealing high levels of engagement and perceived relevance.

Student Interview Findings and Analysis

The qualitative analysis of student interview data revealed four dominant themes: (1) enhanced engagement through personalisation, (2) perceived improvement in vocabulary retention, (3) relevance of contextual learning to real-life experiences, and (4) challenges related to vocabulary difficulty. These themes provide valuable insight into learners' subjective experiences when engaging with AI-generated contextual sentences and complement the quantitative findings from the pre- and post-test results.

Enhanced Engagement Through Personalisation

All ten students reported increased engagement when learning vocabulary through AI-generated contextual sentences. The personalisation of content, particularly sentences aligned with students' interests and daily activities, played a significant role in sustaining attention and motivation. Students indicated that the learning

process felt more interactive and enjoyable compared to traditional vocabulary learning methods such as memorisation of word lists or dictionary-based exercises. This heightened engagement encouraged active participation and willingness to experiment with newly learned vocabulary in classroom tasks.

Several students noted that encountering vocabulary within contexts they could personally relate to reduced anxiety associated with learning unfamiliar words. This finding suggests that personalised contextual input fostered a supportive learning environment that promoted deeper cognitive processing rather than surface-level memorisation.

Perceived Improvement in Vocabulary Retention

Students consistently perceived that their ability to remember vocabulary improved when words were presented in meaningful contexts. Many participants indicated that contextualised exposure helped them recall word meanings more easily during assessments and written tasks. The repetition of vocabulary within varied, personalised sentences further reinforced semantic understanding and facilitated long-term retention.

This perceived improvement aligned with the post-test results, which demonstrated clear gains in definition accuracy and contextual usage. Students reported that seeing how words functioned naturally within sentences enhanced their understanding beyond isolated definitions, enabling them to retrieve vocabulary more effectively when needed.

Relevance of Contextual Learning to Real-Life Experiences

A recurring theme across interviews was the relevance of AI-generated sentences to students' real-life experiences. Students emphasised that vocabulary learning felt more meaningful when examples reflected familiar situations, hobbies, or interests. This relevance supported comprehension by allowing learners to associate new lexical items with existing background knowledge.

The contextual relevance also enhanced students' confidence in applying vocabulary during written tasks. Students indicated that the sentences served as cognitive models, guiding them in constructing their own sentences with appropriate word usage.

Challenges Related to Vocabulary Difficulty

Despite overall positive perceptions, some students expressed difficulty when encountering vocabulary that exceeded their proficiency level. While the contextual sentences were engaging, certain lexical items were perceived as challenging without additional scaffolding or explanation. This challenge highlights the need for adaptive difficulty levels to support learners with varying language proficiency and prevent cognitive overload.

These findings further triangulate quantitative test results and teacher interview data, strengthening the credibility of the study.

Limitations Of the Study

The study is subject to several limitations. The small sample size and single-school setting limit the generalisability of the findings. In addition, the intervention relied primarily on text-based AI input, which may explain the observed gap between written vocabulary mastery and oral fluency. Variations in learners' proficiency levels also suggest that some AI-generated vocabulary exceeded students' linguistic readiness.

Despite these limitations, the findings highlight the pedagogical potential of AI-driven contextual learning when carefully mediated by teachers. Future research should involve larger and more diverse samples, incorporate multimodal features such as audio support, and examine long-term effects on oral vocabulary development. With these refinements, AI-generated contextualisation can offer ESL learners a more effective and practical foundation for vocabulary acquisition.

DISCUSSION AND RECOMMENDATION

Discussion

The use of a mixed-method design allowed for a comprehensive analysis of the intervention's impact. The quantitative results indicated a 100% improvement in specific vocabulary retention tasks, which was strongly supported by the qualitative findings. The convergence of data confirms that AI-generated contextual sentences are a viable pedagogical tool. However, the study also uncovered challenges. Teacher 1 and Teacher 2 noted that the complexity of some vocabulary words was too high for lower-proficiency students, a sentiment echoed by Student 4.

Furthermore, the study's most significant pedagogical finding is the lag in oral production. This aligns with standard distinctions between receptive and productive vocabulary size but highlights a specific limitation of text-based GenAI. The students' hesitation to speak, despite perfect written scores, suggests that the intervention was unimodal (visual/textual). For vocabulary to become truly active, multimodal intervention is required. This suggests that the definition of "Contextual Learning" in the digital age must expand to include audio-visual contexts, potentially utilising Text-to-Speech (TTS) AI features to reinforce phonological acquisition (Wang et al., 2024).

Bridging the Oral-Written Gap

When we talked to teachers, a consistent trend emerged. Students excelled on their written post-tests' errors were rare, and their understanding seemed solid on paper, but when the focus shifted to speaking, they grew hesitant and reserved. They clearly recognised and understood the vocabulary in writing but struggled to use it confidently in conversation. The core of the problem is this: Generative AI, in its current use, is overwhelmingly text-based. It excels at helping students form mental associations between words and their meanings, enriching their passive knowledge. However, it falls short in addressing the phonological loop the system our brains use to process sounds and spoken language. As a result, students can read and write the words, but without opportunities to hear and pronounce them, their oral skills lag behind their written ones. This creates what can be called a "unimodal lag," where one mode of learning outpaces the other. To overcome this, educators should intentionally integrate multimodal AI tools into their lessons. Features like Text-to-Speech and voice-interactive platforms enable students to hear accurate pronunciation and practice speaking in real-time, not just silently read or write. By actively engaging in the auditory and oral aspects of language, students begin to bridge the gap between what they can write and what they can confidently say. This approach doesn't just support better pronunciation; it also builds listening to comprehension and reduces anxiety about speaking, fostering a more holistic language proficiency.

RECOMMENDATIONS

Two major recommendations are proposed for Malaysian ESL educators to optimise AI integration. First, to mitigate cognitive overload, teachers must employ adaptive complexity controls by engineering prompts that explicitly specify proficiency constraints, such as requesting "simple English suitable for Form 4 rural students" rather than relying on generic output. Second, to address the observed lag in oral fluency, a hybridisation of skills approach is essential; educators should immediately follow AI-generated tasks with "Read-Aloud" or "Role-Play" protocols, ensuring that students vocalise the text to activate phonological memory and bridge the gap between written recognition and spoken competence.

Recommendations for ESL Teaching in Malaysia

Let's consider what changes need to happen for AI to truly enhance English language learning in Malaysian classrooms. First, teachers must embrace flexible and responsive prompt engineering. This means adjusting the complexity and style of AI-generated prompts to suit the unique needs of each class. Not every group of students will thrive with the same level of challenge, so having the ability to scale tasks up or down ensures that learners are neither bored nor overwhelmed. Adaptive controls in AI platforms can help teachers personalise instruction, keeping all students engaged and progressing at their own pace. Second, it is essential to integrate language skills, rather than treating them as isolated domains. AI-generated writing should not remain static on the page. Teachers should encourage students to read their writing aloud, participate in dialogues, or perform role-plays using the new vocabulary and structures they have encountered. This active use helps transition words from

passive recognition to genuine, spontaneous expression. When students repeatedly encounter, hear, and use new language in multiple contexts, retention and confidence grow.

Moreover, these strategies should be implemented and refined over extended periods, not just as one-off experiments. Long-term integration allows teachers to track progress, make adjustments, and see how AI can support sustained language development. In this way, AI becomes more than a tool for generating content it evolves into a dynamic partner in the classroom, supporting reading, writing, listening, and speaking in a balanced, interconnected way. With thoughtful application and ongoing evaluation, AI can help Malaysian ESL learners develop not just knowledge, but real communicative competence in English.

CONCLUSION

This mixed-method study confirms that AI-generated contextual sentences are a valuable asset in vocabulary instruction. The triangulated data combining pre-post-test gains with in-depth interviews reveals that this method significantly boosts student engagement, fosters deeper comprehension, and improves retention, particularly in writing. While challenges regarding vocabulary complexity and oral fluency persist, the study suggests that future iterations of this technology should focus on adaptive difficulty levels and integrated speaking opportunities. By implementing these improvements, AI-driven contextual learning can offer ESL students a richer, more practical foundation in vocabulary.

REFERENCES

1. Abdelhalim, S. M., & Alsehibany, R. (2025). Integrating ChatGPT for vocabulary learning and retention: A classroom-based study of Saudi EFL learners.
2. Ahmed, I., Ghafoor, A., Liliuara, D., & Qurrota'Akyuningrum, V. (2025). The impact of AI learning tools on ESL learners' motivation and success: A systematic literature review. English Learning Innovation (englie), 6(1), 134-142.
3. Altaf, A., & Wahida, S. H. (2025). Artificial intelligence-powered vocabulary acquisition: perceptions and performance of university students in Pakistan. Contemporary Journal of Social Science Review, 3(3), 2325-2338.
4. Alih, N. A. C., Raof, A., & Yusof, M. A. M. (2021). Policy change implementation: The case of the CEFR in Malaysian ESL classrooms. Journal of Nusantara Studies, 6(2), 296–317.
5. Alharbi, K., & Khalil, L. (2023). Artificial intelligence (AI) in ESL vocabulary learning: An exploratory study on students and teachers' perspectives. Migration Letters, 20(S12), 1030–1045.
6. Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative research in psychology, 3(2), 77-101.
7. Díaz, B., & Nussbaum, M. (2024). Artificial intelligence for teaching and learning in schools: The need for pedagogical intelligence. Computers & Education, 217, 105071.
8. Elmahdi, O. E. H., Balla, A. A. S., Abdelrady, A. H., Osman, E., & Ahmed, A. O. A. (2025). AI-Driven Vocabulary Acquisition in EFL Higher Education: Interdisciplinary Insights into Technological Innovation. Ethical Challenges, and Equitable Access.
9. Hamedi, W. H. W., Ali, F. D. A., Abdullah, W. Y., Ab Hamid, H., Shuhaimi, N. I. M., & Amir, M. M. (2025). AI as a digital scaffold: An integrative review of Vygotsky's Zone of Proximal Development in modern education. International Journal of Modern Education, 7(26).
10. Johnson,G.B. (2013). Student perception of the flipped classroom. Master thesis, The University of British Columbia.
11. Konyrova, L. (2024). The evolution of language learning: Exploring AI's impact on teaching English as a Second Language. Eurasian Science Review, 2(2), 133–138. https://doi.org/10.63034/esr-42
12. Liu, P. L., & Chen, C. J. (2023). Using an AI-based object detection translation application for English vocabulary learning. Educational Technology & Society, 26(3), 5-20.
13. Majid, N., & Matore, M. E. E. M. (2024). Malaysian ESL teachers' language assessment literacy issues in implementing CEFR-aligned classroom-based assessment: How to deal with them? International Journal of Academic Research in Progressive Education and Development, 13(2).
14. Munshi, A. (2025). A theory of adaptive scaffolding for LLM-based pedagogical agents. arXiv preprint arXiv:2508.01503.

15. Nation, I. S. P. (2022). Learning Vocabulary in Another Language (3rd ed.). Cambridge University Press.
16. Nykyporets, S. S., Pradivlyanny, M. H., Boiko, Y. V., Chopliak, V. V., & Kukharchuk, H. V. (2024). Innovative techniques in vocabulary acquisition for foreign language learning: the impact of artificial intelligence (5): 113-127.
17. Pham, V. P. H., Lian, A., Lian, A., & Barros, S. R. (Eds.). (2025). Implementing AI Tools for Language Teaching and Learning. IGI Global.
18. Sari, N. (2023). The role of artificial intelligence (AI) in developing English language learner's communication skills. Journal on Education, 6(01), 750-757.
19. Schmitt, N. (2023). Vocabulary in Language Teaching (3rd ed.). Cambridge University Press.
20. Schmitt, N. (2021). Vocabulary in Language Teaching (2nd ed.). Cambridge University Press.
21. Vygotsky, L. S. (1978). Mind in Society: The Development of Higher Psychological Processes. Harvard University Press.
22. Wen, Y., Chiu, M., Guo, X., & Wang, Z. (2025). AI-powered vocabulary learning for lower primary school students. British Journal of Educational Technology, 56(2), 734-754.
23. Wang, Y., Li, S., & Zhang, Y. (2024). Empirical assessment of AI-powered tools for vocabulary acquisition in EFL instruction. IEEE Access, 12, 131892–131898.
24. Yasmin Khairani Zakaria, N., Hashim, H., & Azhar Jamaludin, K. (2025). Exploring the Impact of AI on Critical Thinking Development in ESL: A Systematic Literature Review.
25. Zhao, Y. (2023). AI in English Language Teaching: Trends, Challenges, and Opportunities. Journal of Educational Technology, 14(2), 45-60.
26. Zou, M., Luo, J., & Xie, H. (2023). AI-assisted writing and cognitive load in ESL classrooms. Journal of Second Language Writing, 33, 112–127.
27. Zulkefli, N. H., & Ismail, H. H. (2025). The Impact of AI on ESL Writing Autonomy: A Systematic Review. Malaysian Journal of Social Sciences and Humanities (MJSSH), 10(8), e003555-e003555.
28. Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative Research in Psychology, 3(2), 77-101.
29. Nation, I. S. P. (2022). Learning Vocabulary in Another Language. Cambridge University Press.
30. Schmitt, N. (2023). Vocabulary in Language Teaching (3rd ed.). Cambridge University Press.
31. Vygotsky, L. S. (1978). Mind in Society: The Development of Higher Psychological Processes. Harvard University Press.
32. Wang, Y., Li, S., & Zhang, Y. (2024). Empirical assessment of AI-powered tools for vocabulary acquisition in EFL instruction. IEEE Access, 12, 131892–131898.