

The Effectiveness of Using Memrise to Enhance the English Speaking Skills of Thai Students

Muhammad Zhidan Akbar¹, Rita Harisma²

Universitas Muhammadiyah Sumatera Utara^{1,2}

*E-mail: azhidan332@gmail.com

Abstract

Speaking stands out as a core skill for English as a Foreign Language (EFL) learners because it lets them share ideas and actually connect. Still, at Darul Ma'arif School Foundation in Satun, Thailand, I noticed a clear pattern: eighth graders struggled with English speaking. The issues popped up in four main areas—limited vocabulary, shaky pronunciation, slow fluency, and grammar mistakes. To tackle these challenges, I used the Memrise app and set up a quantitative study with a quasi-experimental design. One group of 18 eighth graders—all students in that grade—took a speaking test before and after using Memrise. The test focused on pronunciation, vocabulary, fluency, and accuracy, checked with an analytical rubric. After collecting the scores, I analyzed them using descriptive statistics and a paired-sample t-test. The results told a strong story. Average speaking scores jumped from 10.78 before Memrise to 14.94 after—a boost of 4.16 points. The t-test showed significant improvement, $t(17) = 12.41$, $p < .001$, with a huge effect size (Hedges' $g = 2.79$). Breaking this down, pronunciation improved the most, then accuracy, followed by fluency and vocabulary. All in all, Memrise truly helped these EFL students speak better English. In this Thai school setting, Memrise proved itself as a powerful Mobile-Assisted Language Learning (MALL) tool.

Keywords: Memrise Application, Speaking Skills, Mobile-Assisted Language Learning (Mall), Efl Students, Speaking Development.



Licensees may copy, distribute, display and perform the work and make derivative works and remixes based on it only if they give the author or licensor the credits ([attribution](#)) in the manner specified by these. Licensees may copy, distribute, display, and perform the work and make derivative works and remixes based on it only for [non-commercial](#) purposes.

Introduction

Speaking is a productive skill that reflects learners' ability to use language in real communication, comprising pronunciation, vocabulary, fluency, and accuracy (Harmer, 2015; Nunan, 2015). Pronunciation concerns the clarity of sounds and intonation; vocabulary reflects the range of active words a speaker can readily use; fluency refers to the smoothness of speech with minimal hesitation; and accuracy concerns grammatical correctness (Luoma, 2021). Because these components rely on frequent, meaningful practice, learners in EFL settings, where exposure to the language outside the classroom is limited, tend to develop a largely passive rather than active command of English (Nation, 2001; Swain, 2005).

This pattern is clearly visible, at present, at Darul Ma'arif School Foundation, Satun, Thailand, where preliminary observation and a subsequent teaching internship identified several concrete, ongoing difficulties among eighth-grade students. Many students show low motivation to speak English, largely because they perceive it as a complex language with an alphabetic and phonological system markedly different from Thai; informal interviews revealed confusion in pronouncing letters such as A, B, C, and D as "aab," "bhaa," "chee," and "dehh," indicating first-language phonological interference that affects both intelligibility and speaking confidence (Derakhshan, 2016).

Students also struggle to retain newly learned vocabulary because of insufficient repeated practice, and frequently produce grammatically inaccurate sentences when speaking. These are not isolated or historical issues but the present, day-to-day condition of the class at the time of this study, underscoring the need for a practical, immediately implementable intervention.

Mobile-Assisted Language Learning (MALL) offers a direct response to this situation. Kukulska-Hulme (2012) argues that MALL allows learners to access language material at any time and place, increasing the intensity of practice beyond classroom hours, while Viberg and Grönlund (2013) describe it as an evolution of computer-assisted learning that supports flexible, repeated, and personalised exposure to language input. More recently, Benlaghrissi and Ouahidi (2024) found that MALL-based instruction produced significantly greater gains in EFL speaking performance than conventional classroom teaching, while Dai and Wu (2023) reported that mobile-delivered pronunciation feedback measurably improved learners' articulation over a short intervention period.

This direction is supported by recent studies on Memrise in EFL speaking instruction. Using a similar quasi-experimental pre-test/post-test design, Satrianto, Tusino, and Nugraeni (2025) and Rahmadewi and Nahartini (2025) both reported statistically significant speaking gains after Memrise-based treatment, but both were conducted with Indonesian learners. What remains unaddressed is whether Memrise works as well for learners whose first language creates a more severe type of phonological interference: unlike Indonesian, Thai does not share the Latin alphabet with English, so Thai students must first map unfamiliar letter-sound correspondences before pronunciation gains can even begin. Existing studies also stop short of reporting the fuller statistical evidence, such as normality checks, effect size, and individual-level gain data, needed to judge how robust that effectiveness really is. The present study addresses this gap by testing Memrise with Thai junior high school students in a closely observed, single-class setting, reporting this fuller evidence alongside the pre-test/post-test comparison.

Consistent with this current situation, the research subjects were determined as follows. What: the effectiveness of Memrise in improving speaking skills within a single, intact class, reported only as aggregate data. Who: eighth-grade junior high school students, referred to throughout via anonymous codes rather than personal names. When: the 2025/2026 academic year, over a 28-day teaching-practice period. Where: Darul Ma'arif School Foundation, Satun, Thailand. Why: this was the only eighth-grade class in the school, and it exhibited, in concentrated form, the exact difficulties described above. How: total sampling, since the population was a single, small, academically homogeneous class, all 18 enrolled students were included rather than a randomly drawn subset.

Accordingly, this study aims to examine the effectiveness of the Memrise application as a Mobile-Assisted Language Learning medium in improving the English speaking skills, namely pronunciation, vocabulary, fluency, and accuracy, of these eighth-grade students, with the intention of contributing both to MALL research and to practical classroom strategies for EFL speaking instruction. This aim is guided by the following research question: does the use of the Memrise application significantly improve the English speaking skills of eighth-grade EFL students at Darul Ma'arif School Foundation? This is tested against the following hypotheses:

H0: There is no significant effect of using the Memrise application on students' English speaking skills.

Ha: There is a significant effect of using the Memrise application on students' English speaking skills.

Methods

a. Research Design

This study employed a quasi-experimental design using a one-group pre-test and post-test format, appropriate when random assignment is not feasible but a causal relationship between treatment and outcome is still of interest (Creswell & Creswell, 2022; Shadish, Cook, & Campbell, 2002). This design was adopted because participants formed an intact classroom group in a real

educational setting, consistent with applied linguistics research that prioritises ecological validity over randomisation (Mackey & Gass, 2022). Although no control group was used, comparing students' speaking performance before and after the Memrise intervention allows measurable learning gains to be identified.

b. Population and Sample

The population comprised all eighth-grade students at Darul Ma'arif School Foundation, Satun, Thailand, in the 2025/2026 academic year, which consisted of a single class. Total sampling was applied, involving all 18 students in that class as the research sample; the rationale for this selection is outlined in the Introduction.

c. Research Procedures and Data Collection

Data were collected through three stages: pre-test, treatment, and post-test. Students first completed a speaking pre-test to establish baseline ability (Creswell & Creswell, 2022). The treatment phase then integrated Memrise into speaking instruction for two weeks, following the principle that structured MALL use requires guided rather than incidental integration (Viberg & Grönlund, 2013), and drawing on spaced repetition and retrieval practice as mechanisms for vocabulary activation (Boers, 2022). A post-test, identical in format and difficulty to the pre-test, was then administered under the same conditions to control for instrumentation effects (Ary et al., 2019). All speaking performances were conducted individually and audio-recorded to support objective scoring, so that any improvement could be attributed to the Memrise intervention rather than to external factors.

d. Research Instruments

The instrument was a performance-based speaking test scored with an analytic rubric covering pronunciation, vocabulary, fluency, and accuracy, consistent with Nunan's (2015) view that speaking competence should be evaluated through authentic oral production rather than indirect written measures. The rubric's four-component analytic structure and 1-5 band descriptors were developed in line with established guidance on constructing analytic rubrics for open-ended speaking tasks (Brown, 2020; Vercellotti & McCormick, 2021), which supports the instrument's content validity. Each student performed a structured monologue of approximately 1-2 minutes, scored on a 1-5 scale per component (total range 4-20). Two raters scored each performance, with the final score taken as the mean of the two ratings.

e. Ethical Considerations

Participants' identities are protected throughout this study through the use of anonymous codes (S-1 to S-18) rather than personal names, and no individually identifying information is reported at any point. Data collection took place as part of the researcher's regular teaching placement at the school, with speaking performances recorded solely for scoring purposes.

f. Data Analysis Technique

Data were analysed descriptively (mean, standard deviation) and inferentially. Prior to hypothesis testing, the normality of the pre-test and post-test scores was checked using the Shapiro-Wilk test, appropriate for a sample of this size; because both distributions were found to be normal, a paired-sample *t*-test (Field, 2024) was conducted to determine whether the pre-test/post-test improvement was statistically significant, rather than the non-parametric Wilcoxon signed-rank test. Effect size was additionally calculated using Cohen's *d_z* and the bias-corrected Hedges' *g* (Hedges, 2025) to gauge the practical magnitude of the improvement in speaking scores.

Results and Discussions

1. Results

a. Description of Research Data

This research was conducted at Darul Ma'arif School Foundation, Satun, Thailand, involving 18 eighth-grade students, using a one-group pre-test and post-test design. Speaking ability was assessed before (pre-test) and after (post-test) the Memrise treatment using an analytical

rubric covering pronunciation, vocabulary, fluency, and accuracy, each scored on a 1-5 scale (maximum total score of 20).

b. Comparison of Pre-Test and Post-Test Results

Table 1.
Comparison of Mean Scores by Speaking Component

Speaking Component	Mean Pre-Test	Mean Post-Test	Improvement
Pronunciation	2.72	3.94	1.22
Vocabulary	2.83	3.50	0.67
Fluency	2.67	3.78	1.11
Accuracy	2.56	3.72	1.16
Total Mean	10.78	14.94	4.16

As shown in Table 1, the overall mean score rose from 10.78 in the pre-test to 14.94 in the post-test, an improvement of 4.16 points. Before the intervention, pronunciation ($M = 2.72$) was affected by mispronunciation of sounds that differ from Thai phonological patterns, vocabulary ($M = 2.83$) was limited to basic, frequently repeated words, fluency ($M = 2.67$) was marked by long pauses, and accuracy ($M = 2.56$), the lowest component, reflected frequent errors in basic structures such as to be and the simple present tense. After the two-week Memrise treatment, all four components improved: pronunciation showed the largest gain (1.22 points, rising to 3.94), followed by accuracy (1.16, rising to 3.72), fluency (1.11, rising to 3.78), and vocabulary (0.67, rising to 3.50), the smallest.

c. Treatment Using the Memrise Application

The treatment was conducted for two weeks using Memrise as a MALL medium for teaching speaking, covering daily conversation, self-introduction, classroom expressions, and simple descriptive speaking. Students used Memrise's native-speaker audio feature to improve pronunciation and intonation, while its spaced repetition method supported vocabulary retention and, over time, increased their confidence in speaking English.

d. Normality Test

Before hypothesis testing, the normality of the pre-test and post-test scores was examined using the Shapiro-Wilk test, appropriate for a sample of this size ($n = 18$). As shown in Table 2, both significance values were higher than 0.05 (pre-test $p = .287$; post-test $p = .298$), and the difference scores were likewise normally distributed ($W = 0.922$, $p = .138$); a paired-sample t-test was therefore used rather than the non-parametric Wilcoxon signed-rank test.

Table 2.
Descriptive Statistics and Normality Test

Data	N	Min	Max	Mean	SD	Shapiro-Wilk (W)	Sig.
Pre-Test	18	8	13	10.78	1.52	0.940	0.287
Post-Test	18	11	19	14.94	2.07	0.941	0.298

e. Paired-Sample t-Test and Effect Size Results

Table 3.
Paired Samples Test Results

Mean Diff.	SD	SE Mean	95% CI Lower	95% CI Upper	t	df	Sig. (2-tailed)
4.17	1.43	0.34	3.46	4.88	12.41	17	0.000

As shown in Table 3, the mean difference between post-test and pre-test scores was 4.17 (SD = 1.43, SE = 0.34, 95% CI [3.46, 4.88]), and this difference was statistically significant, $t(17) = 12.41$, $p = .000$, two-tailed; H_a was accepted and H_0 was rejected.

Table 4.
Effect Size

	Point Estimate	95% CI Lower	95% CI Upper
Cohen's dz	2.925	1.839	3.995
Hedges' g	2.794	1.756	3.816

Beyond statistical significance, Table 4 shows that the practical magnitude of this improvement was also very large: Cohen's $d_z = 2.93$, 95% CI [1.84, 4.00], and the bias-corrected Hedges' $g = 2.79$, 95% CI [1.76, 3.82], both well above the conventional threshold of 0.8 for a “large” effect, and with confidence intervals that remain clearly above zero. Taken together, these results indicate that the Memrise application was effective, consistently and substantially, in improving the speaking ability of eighth-grade students at Darul Ma'arif School Foundation, Satun, Thailand.

2. Discussions

The results show that Memrise had a positive effect on all four speaking components, but the size of that effect was not uniform, and this pattern is worth explaining rather than simply restating. Pronunciation improved the most (1.22 points), consistent with the mechanism Memrise actually provides: repeated, on-demand exposure to native-speaker audio gives learners an explicit, correct model to imitate, which directly targets the first-language phonological interference identified in the Introduction (Derakhshan, 2016) and aligns with evidence that mobile-delivered pronunciation feedback produces measurable articulation gains even over short intervention periods (Dai & Wu, 2023; Ngo et al., 2023). Accuracy showed the second-largest gain (1.16 points), plausibly because Memrise's sentence-pattern drills function as a form of retrieval practice: each attempt to reproduce a grammatical pattern from memory, rather than simply re-reading it, is understood to strengthen retention more effectively than passive review (Boers, 2022).

Vocabulary showed the smallest gain (0.67 points), which at first appears to sit awkwardly alongside Memrise's central use of spaced repetition for vocabulary building. A likely explanation is that spaced-repetition systems are designed to distribute retrieval over long intervals, so their benefit compounds over weeks or months; a two-week treatment window may simply be too short for the spacing effect to reach its full strength, even though it is enough time for more immediate gains such as pronunciation modelling to appear. This suggests that vocabulary gains in particular may continue to rise with continued use beyond the period measured here, rather than indicating that Memrise is ineffective for vocabulary. The fluency gain (1.11 points), meanwhile, is consistent with Swain's (2005) Output Hypothesis, under which repeated pushed production gradually reduces the cognitive effort needed to retrieve words and structures in real time, producing smoother speech with fewer pauses.

These findings should be read with some caution regarding alternative explanations. Because this study used a one-group design without a control group, part of the improvement could plausibly reflect factors other than Memrise itself, such as a novelty effect, a practice or testing effect from performing the speaking task twice, or increased teacher attention during the treatment period. The relatively small standard deviation of the gain scores ($SD = 1.43$) next to a mean gain of 4.17, together with the very large effect size, makes it unlikely that these factors alone account for the full improvement, but a true experimental design with a control group would be needed to rule them out with more confidence.

These results align with Satrianto, Tusino, and Nugraeni (2025) and Rahmadewi and Nahartini (2025), who similarly found Memrise effective in improving EFL students' speaking ability, and extend that evidence to Thai learners facing a distinct, alphabet-level form of phonological interference. Taken together, the findings support Memrise as an effective Mobile-Assisted Language Learning medium for improving pronunciation, vocabulary, fluency, and accuracy in EFL speaking, while highlighting pronunciation and accuracy as the components that benefit most within a short treatment window.

Conclusion

This quasi-experimental study found that the Memrise application was effective in improving the English speaking skills of eighth-grade students at Darul Ma'arif School Foundation. Students' overall mean speaking score increased from

10.78 to 14.94, a gain of 4.16 points, and the paired-sample t-test, $t(17) = 12.41$, $p = .000$, confirmed that this improvement was statistically significant and very large in magnitude (Hedges' $g = 2.79$), so that H_0 was rejected and H_a was accepted. Among the four components, pronunciation showed the largest gain (1.22 points), followed by accuracy (1.16), fluency (1.11), and vocabulary (0.67); as discussed above, this pattern is consistent with native-speaker audio and retrieval-based drills most directly supporting pronunciation and grammatical accuracy in the short term, while vocabulary gains from spaced repetition typically compound over a longer period (Nation, 2001; Schmitt, 2020).

These findings support Memrise as an effective Mobile-Assisted Language Learning tool for EFL speaking instruction. English teachers are encouraged to integrate MALL applications such as Memrise as a supplementary medium to increase students' opportunities for autonomous speaking practice, particularly in EFL contexts where exposure to authentic English input is limited.

References

- Alfina Rahmadewi, D. N. (2025). The impact of Memrise application on students' speaking skills (a quasi-experimental study at Islamic junior high school). *Journal on Language Teaching and Learning, Linguistics and Literature*, 8486–8499.
- Ary, D., Jacobs, L., Irvine, C., & Walker, D. (2019). *Introduction to research in education* (10th ed.). Cengage Learning.
- Benlaghrissi, H., & Ouahidi, L. M. (2024). The impact of mobile-assisted project-based learning on developing EFL students' speaking skills. *Smart Learning Environments*, 11(1), 18. <https://doi.org/10.1186/s40561-024-00303-y>
- Boers, F. (2022). Evaluating second language vocabulary and grammar instruction. *Language Teaching Research*, 26(3), 321–335.
- Brown, H. D. (2020). *Principles of language learning and teaching* (7th ed.). Pearson.
- Creswell, J. W., & Creswell, J. D. (2022). *Research design: Qualitative, quantitative, and mixed methods approaches* (6th ed.). Sage.
- Dai, Y., & Wu, Z. (2023). Mobile-assisted pronunciation learning with feedback from peers and/or automatic speech recognition: A mixed-methods study. *Computer Assisted Language Learning*, 36(5–6), 861–884. <https://doi.org/10.1080/09588221.2021.1952272>

- Derakhshan, A., Khalili, A. N., & Beheshti, F. (2016). Developing EFL learner's speaking ability, accuracy and fluency. *English Language and Literature Studies*, 6(2), 177–186. <https://doi.org/10.5539/ells.v6n2p177>
- Field, A. (2024). *Discovering statistics using IBM SPSS statistics* (6th ed.). Sage.
- Harmer, J. (2015). *How to teach English* (2nd ed.). Longman.
- Hedges, L. V. (2025). Effect sizes for experimental research. *British Journal of Mathematical and Statistical Psychology*. <https://doi.org/10.1111/bmsp.12389>
- Kukulska-Hulme, A. (2012). Mobile-assisted language learning. In C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics*. Blackwell Publishing.
- Luoma, S. (2021). *Assessing speaking* (2nd ed.). Cambridge University Press.
- Mackey, A., & Gass, S. (2022). *Second language research: Methodology and design* (3rd ed.). Routledge.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge University Press.
- Nunan, D. (2015). *Teaching English to speakers of other languages: An introduction*. Routledge.
- Richards, J. C., & Renandya, W. A. (2002). *Methodology in language teaching*. Cambridge University Press.
- Satrianto, F., Tusino, T., & Nugraeni, I. I. (2025). The effectiveness of using Memrise application in improving students' speaking skills. *JiIP - Jurnal Ilmiah Ilmu Pendidikan*, 8(8), 9141–9149. <https://doi.org/10.54371/jiip.v8i8.8850>
- Schmitt, N. (2020). *Vocabulary in language teaching* (2nd ed.). Cambridge University Press.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin.
- Swain, M. (2005). The output hypothesis: Theory and research. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 471–483).
- Thi-Nhu Ngo, T., Hao-Jan Chen, H., & Kuo-Wei Lai, K. (2023). The effectiveness of automatic speech recognition in ESL/EFL pronunciation: A meta-analysis. *ReCALL*, 36(1), 1–18. <https://doi.org/10.1017/s0958344023000113>
- Vercellotti, M. L., & McCormick, D. E. (2021). Constructing analytic rubrics for assessing open-ended tasks in the language classroom. *TESL-EJ*, 24(4), 1–19.
- Viberg, O., & Grönlund, Å. (2013). Systematising the field of mobile assisted language learning. *International Journal of Mobile and Blended Learning (IJMBL)*, 5(4), 72–90.

Acknowledgment

The author would like to express sincere gratitude to Darul Ma'arif School Foundation, Satun, Thailand, and its eighth-grade students for their participation and cooperation throughout this research. Special thanks are also extended to the supervisor, Rita Harisma, S.Pd., M.Hum., for her invaluable guidance, feedback, and support throughout the completion of this study.