Journal of English Language Teaching and Learning (JETLE)
Vol 8, No 1, 2026, Page 1-11
● ISSN 2686-665X

The effects of flashcards-assisted substitution drills on EFL students' vocabulary mastery: A quasi-experimental study in a junior secondary school

Lambreta Alya Rani Fitria[1], Fenty Andriani[2]

[1,2]Universitas Islam Negeri Kiai Ageng Muhammad Besari Ponorogo
[1]lambreta.alya@gmail.com, *[2]fentyandriani@uinponorogo.ac.id

Keywords:
vocabulary mastery, substitution drill technique, flashcards

A B S T R A C T

Vocabulary mastery is a key element in English as a Foreign Language (EFL) learning, underpinning all four language skills. However, students at SMP Negeri 2 Babadan Ponorogo still face difficulties in retaining and applying vocabulary due to conventional, rote-memorization-based instruction. This study aimed to investigate the effectiveness of the Substitution Drill Technique using flashcards in improving students' vocabulary mastery. A quasi-experimental design with a non-equivalent control group was employed. The students of SMP Negeri 2 Babadan Ponorogo in academic year 2025/2026. Based on the 72 Year 8 students, 48 students from two full classes were purposively selected according to their availability within the timetable. Class VIII A (24 students) was selected as the experimental group, whilst Class VIII B (24 students) was selected as the control group. The experimental group was given Substitution Drill Technique using flashcards and the control group was given translation-based vocabulary technique using infographic media. Pre-test and post-test scores were analyzed using IBM SPSS Statistics 21. In the experimental group the mean score improved from 47.12 to 64.46, and in the control group it went from 49.71 to 58.58. The mean N-Gain scores were 32.36 and 13.11 respectively. An independent-samples t-test comparing N-Gain scores between the experimental and control groups showed a statistically significant difference ($t = 2.547$, $df = 46$, $p = .014$). Because this is a quasi-experimental design and single school, the finding should be read as evidence that there was a significant difference between groups and not as proof of causality.

1. INTRODUCTION

Vocabulary mastery is one of the most important aspects of the English as a Foreign Language (EFL) learning process, as it serves as the foundation upon which all four language skills—listening, speaking, reading, and writing—are built. Even a highly competent learner of grammar needs a sufficient stock of vocabulary items in order to comprehend and express messages successfully (Hartini & Suri Ardini, 2024). Learners who do not have adequate vocabulary knowledge experience significant problems in expressing themselves clearly, both orally and in writing, and in understanding what they hear and read. Vocabulary knowledge extends beyond knowledge of word meanings to include knowledge of word form and appropriate use in different communicative contexts, such as grammatical patterns, collocations, and register-appropriate variations (Hartini & Suri Ardini, 2024). Despite this significance,

vocabulary learning might not be given the priority that would seem to be warranted when other language skills, such as grammar, reading comprehension, or other aspects of the language, are emphasized.

A continuing problem in EFL vocabulary learning is the gap between receptive and productive vocabulary knowledge. Students may recognize words when reading or listening but have difficulty retrieving and using them when speaking or writing. This gap may occur when learners have limited opportunities to retrieve and use newly learned words in meaningful language activities (Lei & Reynolds, 2022). When learners engage in passive learning activities like translating or memorizing a word list, they may see but fail to use enough words actively. Therefore, vocabulary instruction needs to go beyond recognition and provide systematic opportunities for learners to retrieve and produce vocabulary in context.

The majority of EFL classes, specifically at lower middle school level, employ a variety of vocabulary teaching practices, including translation, textbook exercises, repetitive drills and visual learning activities. Previous research in the context of EFL in Indonesia has reported the use of teacher-led and practice-based activities, although teaching practices may vary across different classroom contexts (Suwarno et al., 2023). However, effective vocabulary learning requires not only deliberate exposure to new vocabulary, but also opportunities for learners to actively retrieve and apply it (Hartini & Suri Ardini, 2024; Lei & Reynolds, 2022). Consequently, teaching that relies heavily on memorization, translation, or textbook exercises may provide only limited opportunities for productive vocabulary use. Teaching activities that combine repeated exposure with active recall and structured use can create a learning environment that better supports vocabulary development.

Flashcards are a learning material that can support vocabulary learning by providing visual guidance to link word forms with their meanings. Repeated exposure and recall exercises using physical or digital flashcards can reinforce vocabulary learning. Research by Azahra et al. (2024), Hasanah et al. (2024), Wulandari (2025) indicates that flashcards can support vocabulary learning. However, flashcards primarily provide visual support and do not always require learners to use the target vocabulary at the sentence level. Therefore, combining flashcards with activities that require learners to recall and use the target words can extend their instructional value from recognition and recall to productive use.

The Substitution Drill technique is a systematic and structured approach that can complement the use of flashcards by providing opportunities for controlled language production. This technique is generally used to practice sentence structures through repetition and substitution, which involves learners replacing specific elements in a sentence with elements of the target language (Saefudin & Bakhtiar, 2024). Suwarno et al. (2023) reported that repetition, controlled practice, and drilling are among the techniques used in the Audio-Lingual Method to support language learning. Previous research has examined substitution exercises across various areas of language learning. Hidayati and Nurhaera (2023) investigated substitution exercises in the teaching of the present progressive tense, while Lutfiyah et al. (2022) explored their application in the teaching of the simple present tense. Saefudin and Bakhtiar (2024) reported a significant effect of substitution exercises on students' English-speaking skills. These studies indicate that substitution exercises can provide learners with structured opportunities to repeatedly recall and produce linguistic elements. However, previous studies have primarily focused on grammar or speaking skills, rather than vocabulary acquisition.

The combination of flashcards and substitution drills is based on the complementary functions of these two teaching components. Flashcards can provide visual cues to support vocabulary recognition and recall, while substitution drills may require learners to recall the target vocabulary and use it in repetitive sentence patterns. Previous research has generally examined these approaches separately: studies on substitution exercises have primarily focused on grammar or speaking performance (Hidayati et al., 2023; Lutfiyah et al., 2022; Saefudin & Bakhtiar, 2024), and research on flashcards has largely centered on vocabulary acquisition and retention (Azahra et al., 2024; Hasanah et al., 2024; Xodabande et al., 2022). Consequently, research examining the combined use of flashcards and substitution drills to support vocabulary knowledge in terms of form, meaning and use, particularly amongst Indonesian lower middle school students learning English as a foreign language (EFL), remains limited. This study addresses this gap by investigating this blended teaching approach within the context of lower secondary schools in Indonesia.

This study was conducted at SMPN 2 Babadan, Ponorogo, where the researchers' initial observations and informal interviews with the English teacher indicated that vocabulary was one of the areas in which students experienced difficulties, particularly in retaining and appropriately using vocabulary in speaking and writing activities. These findings were obtained from the researchers' initial observations and informal interviews rather than from previous research. The initial observations also indicated that vocabulary learning tended to involve textbook-oriented activities, book work, and test-type exercises, with limited opportunities for students to use vocabulary in meaningful situations. These classroom conditions are relevant to broader challenges reported in rural and non-technological EFL settings, where teachers may need to adapt instructional strategies to available resources (Candrawati & Purbani, 2025). Hence, the use of visual support with flashcards and structured production using substitution drills were identified as relevant in addressing the identified classroom needs.

Based on the identified classroom problem and the research gap concerning the limited investigation of the combined use of flashcards and substitution drills, this study aims to examine the effectiveness of the combined approach in improving eighth-grade students' vocabulary mastery at SMPN 2 Babadan, Ponorogo. The study focuses on vocabulary knowledge in terms of form, meaning, and use and seeks to provide empirical evidence from an Indonesian junior high school EFL context.

2. METHOD

This study employed a quantitative approach using a quasi-experimental design with a non-equivalent control group. The population comprised 72 eighth-grade students in three intact classes (VIII A, VIII B, and VIII C) at SMPN 2 Babadan Ponorogo during the 2025/2026 academic year. The first two classes, VIII A and VIII B, were purposively selected based on their availability in the instructional schedule. Class VIII C was excluded because the instructional time for the class was not sufficient for them to participate. The final sample size was 48 students, representing 24 students in each selected class. VIII A was assigned as the experimental group and VIII B as the control group. Pre-test scores were then used to check the comparability of the baseline. The Substitution Drill Technique using flashcards was given to the experimental group while the control group was given vocabulary learning based on translation technique using infographic media. The independent variable was the Substitution Drill Technique combined with flashcards, while the dependent variable was students' vocabulary mastery, operationalized in terms of vocabulary form, meaning, and use.

The test consisted of multiple-choice, cloze, and word-form items. Content validity was determined through expert judgment by two lecturers in English language teaching with relevant expertise in vocabulary teaching and test design, following principles of language assessment concerning the appropriateness and quality of classroom assessment. Items were evaluated by the experts using a content-validity checklist that included curriculum relevance, clarity of instructions, appropriateness of cognitive level, and representation of the vocabulary dimensions assessed. Their feedback was used to revise the instrument before the try-out. Pearson's bivariate correlation was thereafter tested for 24 students at the 0.05 level of significance in order to evaluate the empirical item validity, following the principles of educational evaluation (Putri et al., 2024). With df = 22 and r table = 0.404, all 30 items met the validity criterion and were retained for the study.

The internal consistency of the vocabulary test was examined using Cronbach's alpha based on responses from 24 students participating in the try-out (Putri et al., 2024). Both the pre-test and post-test tools had high internal consistency with Cronbach's alpha of 0.895 and 0.934, respectively.

Due to the non-equivalent control group design, pre-test scores were used to check for baseline comparability. The experimental group obtained a mean score of 47.125, whereas the control group obtained 49.70833. The Shapiro–Wilk test showed that the experimental-group scores were normally distributed ($p = 0.173$), whereas the control-group scores were not ($p = 0.044$). Therefore, the Mann–Whitney U test was used to compare the initial vocabulary mastery of the two groups. The analysis showed that there was no statistically significant difference between the groups before the treatment ($U = 260.000$, $p = 0.559$). Therefore, no statistically significant difference in initial vocabulary mastery was found between the two groups before the treatment.

The treatment was conducted over four meetings of 2 × 40 minutes. The first meeting in the experimental group was started by administering the pre-test, followed by introducing words related to invitation with flashcards, invitation-related word pronunciation, meaning identification, sentence completion, and grouping words in the experimental group. The second meeting included a review of the target vocabulary using the flashcard guessing activity and substitution drills using invitation sentence patterns. Students practiced replacing selected words or phrases using flashcard cues through whole-class, pair, and small-group activities. The third meeting focused on advertisement-related vocabulary, followed by guided and group substitution drills and the production of simple advertisement sentences, which were shared for feedback. In the fourth meeting, students reviewed invitation and advertisement vocabulary and sentence patterns through flashcard activities, completed individual sentence-writing practice, and then took the post-test.

The control group was given the same amount of learning materials and the same amount of instructional time as the experimental group but not the flashcard and substitution-drill treatment. In contrast, during four meetings, the instruction was carried out using translation, teacher explanation, infographic media, and written exercises. This allowed the study to compare the vocabulary outcomes of students receiving the combined instructional treatment with those receiving the conventional instructional approach.

The data were analyzed using IBM SPSS Statistics 21. Descriptive statistics were used to summarize the pre-test and post-test scores. The Shapiro–Wilk test was used to examine normality, and the Mann–Whitney U test was used to examine baseline comparability because the control-group pre-test data were not normally distributed. N-Gain scores were calculated

from the pre-test and post-test scores. After the normality and homogeneity assumptions were examined, an independent-samples t-test was used to compare N-Gain scores between the experimental and control groups.

3. RESULTS AND DISCUSSION

Results

Pre-Test and Post-Test Scores of the Experimental Class

Table 1 presents the pre-test and post-test vocabulary scores of the experimental class (VIII A), comprising 24 students who received the Substitution Drill Technique combined with flashcards.

Table 1. Pre-Test and Post-Test Scores of the Experimental Class (VIII A)

No.	Student	Pre-Test	Post-Test
1	S1	60	77
2	S2	47	57
3	S3	44	64
4	S4	57	84
5	S5	40	47
6	S6	44	67
7	S7	54	77
8	S8	40	60
9	S9	40	50
10	S10	64	64
11	S11	44	47
12	S12	54	80
13	S13	44	74
14	S14	40	67
15	S15	50	80
16	S16	24	60
17	S17	37	67
18	S18	40	77
19	S19	50	47
20	S20	44	47
21	S21	40	50
22	S22	57	80
23	S23	60	67
24	S24	57	57
Total		1131	1547
Mean		47.12	64.46

According to Table 1, the total pre-test scores of the experimental class were 1,131 and the mean was 47.12, meaning that the experimental class had overall low level of vocabulary mastery before the intervention, ranging from the lowest of 24 to the highest of 64. After the application of the Substitution Drill Technique with flashcards, the post-test total rose to 1,547 with a mean of 64.46 — an improvement of 17.34 points — with the minimum rising to 47 and maximum to 84. The experimental group showed an increase in vocabulary scores following the instructional treatment, with a mean improvement of 17.34 points.

Pre-Test and Post-Test Scores of the Control Class

Table 2 presents the pre-test and post-test vocabulary scores of the control class (VIII B), comprising 24 students who received conventional translation-based instruction without substitution drills or flashcards.

Table 2. Pre-Test and Post-Test Scores of the Control Class (VIII B)

No.	Student	Pre-Test	Post-Test
1	S1	57	64
2	S2	70	50
3	S3	50	47
4	S4	54	67
5	S5	44	70
6	S6	44	64
7	S7	64	54
8	S8	34	64
9	S9	40	64
10	S10	44	54
11	S11	50	70
12	S12	50	60
13	S13	67	74
14	S14	40	54
15	S15	40	50
16	S16	44	57
17	S17	60	64
18	S18	44	60
19	S19	40	60
20	S20	37	47
21	S21	40	60
22	S22	60	54
23	S23	70	54
24	S24	50	44
Total		1193	1406
Mean		49.71	58.58

Table 2 shows that the control class achieved a pre-test total of 1,193 with a mean of 49.71 (lowest: 34, highest: 70), while the post-test total rose to 1,406 with a mean of 58.58 (lowest: 44, highest: 74), representing an improvement of 8.87 points.

Descriptive Statistics of Pre-Test

Table 3. Descriptive Statistics of Pre-Test Scores

Descriptive Statistic	Experimental Class (VIII A)	Control Class (VIII B)
Mean	47.12	49.71
Minimum	24	34
Maximum	64	70

The pre-test was administered to both the experimental and control groups to determine their initial vocabulary mastery before the treatment. The experimental group consisted of 24 students with scores ranging from 24 to 64, while the control group consisted of 24 students with scores ranging from 34 to 70. The mean pre-test score was 47.12 for the experimental group and 49.71 for the control group. Therefore, the descriptive mean scores of the two groups were not very different prior to the treatment.

The normality of the pre-test scores was examined using the Shapiro–Wilk test. The experimental group obtained a significance value of 0.173, indicating normally distributed data, whereas the control group obtained 0.044, indicating that the data did not meet the normality assumption. Therefore, the Mann–Whitney U test was used to examine the difference between the two groups.

Normality Test and Mann-Whitney U Test of Pre-Test

Table 4. Shapiro-Wilk Normality Test – Pre-Test

Class	Shapiro-Wilk Statistic	f	ig.
Experimental (VIII A)	.941	4	173
Control (VIII B)	.915	4	044

The Shapiro-Wilk normality test (Table 4) showed that the experimental class pre-test data were normally distributed ($p = 0.173 > 0.05$), while the control class data were not normally distributed ($p = 0.044 < 0.05$). Since the normality assumption was violated by the control group, the Mann-Whitney U Test was applied as a non-parametric alternative.

Table 5. Mann-Whitney U Test Result – Pre-Test

Test Statistic	Value
Mann-Whitney U	260.000
Wilcoxon W	560.000
Z	-.584
Asymp. Sig. (2-tailed)	0.559

The Mann–Whitney U test yielded a significance value of 0.559, which was greater than 0.05. Therefore, no statistically significant difference in pre-test vocabulary scores was detected between the experimental and control groups. This finding suggests that no statistically significant difference between the groups was found prior to the treatment.

Descriptive Statistics of Post-Test

Table 6. Descriptive Statistics of Post-Test Scores

Descriptive Statistic	Experimental Class (VIII A)	Control Class (VIII B)
Mean	64.46	58.58
Minimum	47	44
Maximum	84	74

After the instructional treatment, both groups completed the post-test to measure their vocabulary mastery. The experimental group had scores ranging from 47 to 84 with a mean score of 64.46 while the control group had scores ranging from 44 to 74 with a mean score of 58.58. Compared with their respective pre-test scores, the experimental group showed a mean increase of 17.34 points, while the control group showed an increase of 8.87 points.

The experimental group obtained a higher post-test mean score than the control group. This difference is described here as a descriptive difference, while the inferential comparison reported in this study focuses on the N-Gain scores.

N-Gain Analysis

Table 7. Descriptive Statistics of N-Gain Scores

N-Gain Statistic	Experimental Class (VIII A)	Control Class (VIII B)
N (Valid)	24	24
Mean	32.36	13.11
Minimum	-6.00	-66.67
Maximum	62.79	46.43

N-Gain scores were calculated to describe the improvement in students' vocabulary scores from the pre-test to the post-test. The experimental group obtained a mean N-Gain score of

32.36, while the control group obtained a mean of 13.11. The experimental group showed greater descriptive improvement in N-Gain scores than the control group.

Normality and Homogeneity Tests of N-Gain

Table 8. Shapiro-Wilk Normality Test – N-Gain

Class	Shapiro-Wilk Statistic	df	Sig.
Experimental (VIII A)	.918	24	.053
Control (VIII B)	.968	24	.627

The N-Gain data were normally distributed in both groups, with Shapiro–Wilk significance values of 0.053 for the experimental group and 0.627 for the control group.

Because both N-Gain significance values were greater than 0.05, the homogeneity of variance was examined using Levene's test before the independent-samples t-test.

Table 9. Levene's Homogeneity Test – N-Gain

Basis	Levene Statistic	df1	df2	Sig.
Based on Mean	.226	1	46	.637
Based on Median	.017	1	46	.896
Based on Median (adj. df)	.017	1	35.073	.896
Based on Trimmed Mean	.059	1	46	.810

Based on the mean, Levene's test yielded a significance value of 0.637, which was greater than 0.05. Therefore, the variances of the N-Gain scores between the experimental and control groups were considered homogeneous, and the equal-variances-assumed result was used for the independent-samples t-test.

Independent Sample T-Test (N-Gain)

Examination of the assumptions of normality and homogeneity led to an independent-samples t-test for the comparison of N-Gain scores between the experimental and control groups.

Table 10. Result of Independent Sample T-Test of N-Gain

Variance	F	Sig.	t	df	Sig. (2-tailed)	Mean Diff.	Std. Error
Equal variances assumed	.226	.637	2.547	46	.014	.19249	.07556
Equal variances not assumed			2.547	42.886	.015	.19249	.07556

The independent-samples t-test for N-Gain scores resulted in a t-value of 2.547, with 46 degrees of freedom, and a two-tailed significance value of 0.014. Levene's test indicated homogeneous variances ($F = 0.226$, $p = 0.637$); therefore, the equal-variances-assumed result was used. Since $p = 0.014$ was less than 0.05, the difference in N-Gain scores between the experimental and control groups was statistically significant. The experimental group also had a higher mean N-Gain score (32.36) than the control group (13.11).

Discussion

The main finding of this study is that the experimental group achieved a significantly higher mean N-Gain score than the control group. The N-Gain scores for both groups showed improvement between pre- and post-test, but the difference between the groups was statistically significant, $t(46) = 2.547$, $p = .014$. This result indicates that the two teaching conditions were associated with significantly different levels of vocabulary improvement during the instructional period.

The pattern observed can be explained in comparison to the complementary skills of flashcards and substitution drills. While flashcards offer visual support for form recognition and

retrieval and word meaning, substitution drills offer repeated opportunities to retrieve and produce target vocabulary in a structured sentence pattern. These activities may have given a sequence from visual identification and retrieval to repeated use of target vocabulary at the sentence level. This is a possible explanation of the pattern observed but is not evidence of a causal mechanism.

The findings can be related to previous studies that have reported positive results from the use of flashcards, drilling, and substitution drills in language learning. Muhdar and Said (2024), for example, reported the effectiveness of drilling techniques in improving students' vocabulary mastery. Azahra et al. (2024) reported an increase in students' vocabulary scores following the use of flashcards as teaching aids. Laily and Febrianingrum (2023) reported improvement in vocabulary achievement following drilling, while Merdani et al. (2024) also reported improved vocabulary achievement following drilling activities. Saefudin and Bakhtiar (2024) reported significant improvement in English speaking skills following the use of substitution drills. These findings provide relevant evidence for the potential value of visual materials and systematic drilling in language learning. However, the previous studies differed from the present study in participants, instructional context, treatment duration, research design, and/or learning outcomes. Therefore, the present study aimed to investigate the integration of flashcards and substitution drills in vocabulary learning among junior high school EFL students in Indonesia. The statistically significant difference in N-Gain scores obtained in the present study ($p = .014$) should be seen in the specific characteristics of the present research context.

The possible contribution of the combined technique can further be discussed through relevant learning theories. According to Dual Coding Theory, information presented in both verbal and nonverbal codes can aid learning and memory (Li et al., 2022). In the present study, the target vocabulary was presented visually through flashcards and practiced verbally through repeated sentence patterns in the substitution drills. Some of the substitution-drill activities used also had characteristics of the Audio-Lingual Method, such as repetition, controlled practice, and corrective feedback (Suwarno et al., 2023). In addition, embedding target words in sentence contexts may provide opportunities to connect vocabulary with contextual use (Çoban Sural & Yaşar Sağlık, 2024). These theoretical perspectives are offered as possible explanations for the learning gains that occurred in the experimental group but are not evidence of the mechanism that caused the statistically significant difference between the experimental and control groups.

The statement concerning students' engagement should be interpreted cautiously. The present study did not use a quantitative engagement instrument; therefore, engagement cannot be established as a measured outcome of the treatment. The activities carried out in the classroom might have offered opportunities for active participation, but this is an observation-based possibility, not a proven effect.

Regarding the research question, the experimental group demonstrated greater N-Gain improvement than the control group, and the independent-samples t-test showed a statistically significant difference in N-Gain scores, $t(46) = 2.547$, $p = .014$. The results showed that there was a significant difference between the vocabulary gains of the two instructional conditions over the course of the study. The finding is presented as an observed between-group difference and is not a proof of causality since it was a quasi-experimental design with two intact classes in one school. There have also been other contextual or group-level factors that could have influenced the results of the observed differences.

The results have implications for vocabulary teaching. Flashcards and substitution drills may be used as reinforcers and to provide additional visual support and repetition for target

words when these are needed in the classroom. However, teachers should adapt the activities to the learners' level, available instructional time, and learning objectives. Thus, the results should be interpreted with the following limitations. This study took place in one school, with two intact classes, and a relatively short four meeting treatment period, all of which restrict generalizability. The study also did not employ a quantitative measure of student engagement. Larger and more diverse samples, longer treatment duration and additional measures (such as classroom observation or engagement instruments) should be included in future studies.

4. CONCLUSION

The results show that students who were taught by the experimental method made more gains in vocabulary learning than students who were taught by the conventional method. The N-Gain analysis showed a statistically significant difference between the groups, with the experimental group obtaining a higher mean N-Gain score than the control group. The Substitution Drill Technique thus may be a useful supplement to the vocabulary instruction using the flashcards because it is supplemented by visuals and gives opportunities to retrieve and produce vocabulary. The results were obtained from a quasi-experimental design with two intact classes in one school, and a four-meeting treatment duration, however, and should be interpreted within these limitations. Thus, the findings suggest that there is a significant difference between the two instructional conditions in this study, but do not provide clear evidence of causal effects.

REFERENCES

Azahra, N. H. N., Isnaini, S., Sudarmaji, I., Nuraini, H., Forsia, L., & Nirmala, N. (2024). implementing the use of flashcards to improve students' vocabulary mastery; classroom action research. *Jurnal Basicedu*, *8*(1), 225–232. https://doi.org/10.31004/basicedu.v8i1.6896

Candrawati, A. C., & Purbani, W. (2025). Teaching English In rural settings: A systematic review of challenges and strategies in non-technological classrooms. *English Language Teaching Educational Journal*, *8*(2), 66–77. https://doi.org/10.12928/eltej.v8i2.14081

Çoban Sural, Ü., & Yaşar Sağlık, Z. (2024). Next-generation learning experience: The effect of technology-assisted and printed context-based vocabulary activities on vocabulary knowledge and learning motivation. *Education and Information Technologies*, *29*(16), 20823–20853. https://doi.org/10.1007/s10639-024-12688-x

Hartini, L. W., & Suri Ardini, A. (2024). Relationship between Indonesian EFL learners' vocabulary knowledge and English competencies. *Journal of English and Education (JEE)*, 55–66. https://doi.org/10.20885/jee.v10i1.33651

Hasanah, U., Oktaviana, F., & Fitriani, D. (2024). The use of electronic flashcards: Enhancing vocabulary mastery through classroom action research. *Jurnal Ilmiah Bina Bahasa*, *17*(2).

Hidayati, Yunus H, I., & Nurhaera. (2023). Teaching present progressive tense to the second-year students of junior high school through substitution drill at SMP 6 Tolitoli. *Jurnal Madako Education*, *9*(2), 6–11.

Laily, U., & Febrianingrum, L. (2023). The effect of drilling routine tasks on students' vocabulary mastery of beginner English class (BEC) at Islamic boarding school Pamekasan. *PANYONARA: Journal of English Education*, *5*(1), 30–40. https://doi.org/10.19105/panyonara.v5i1.6974

Lei, Y., & Reynolds, B. L. (2022). Learning English vocabulary from word cards: A research synthesis. *Frontiers in Psychology*, *13*, 984211. https://doi.org/10.3389/fpsyg.2022.984211

Li, W., Yu, J., Zhang, Z., & Liu, X. (2022). Dual coding or cognitive load? exploring the effect of multimodal input on English as a foreign language learners' vocabulary learning. *Frontiers in Psychology*, *13*, 834706. https://doi.org/10.3389/fpsyg.2022.834706

Lutfiyah, L., Sadikin, I. S., Fuady, F., Hidayat, R. K., & Frandikta, Y. (2022). The Implementation of substitution drill in teaching simple present tense. *Pedagonal : Jurnal Ilmiah Pendidikan*, *6*(2), 221–227. https://doi.org/10.55215/pedagonal.v6i2.5657

Merdani, N. L., Dewi, K. S., & Mahendrayana, G. (2024). The use of drilling techniques integrated with technology to improve students' vocabulary mastery. *Journal of Educational Study (JoES)*, *4*(2), 56–69.

Muhdar, F., & Said, D. R. (2024). The effectiveness of drilling technique on students' vocabulary mastery. *The 2nd International Conference on Language Teaching*, 343–354.

Putri, A. M. K., Kristanto, B., & Soali, M. (2024). Examining validity and reliability of general English course assessments in higher education. *Fonologi : Jurnal Ilmuan Bahasa Dan Sastra Inggris*, *2*(3), 325–332. https://doi.org/10.61132/fonologi.v2i3.1107

Saefudin, D. P., & Bakhtiar, A. (2024). The effect of substitution drill technique on students' English-speaking skill. *Journal of English Language Learning (JELL)*, *8*(1), 471–478.

Suwarno, T. S., Retnowati, N., & Sundari, H. (2023). Exploring teacher's implementation of audio-lingual method, challenges, and techniques for improving student's vocabulary mastery. *ENGLISH JOURNAL*, *17*(1), 23–31. https://doi.org/10.32832/english.v17i1.8891

Wulandari, E. M. (2025). The effectiveness of using flashcards on teaching English vocabulary. *Journal Of English Education And Technology*, *5*(4), 324–338.

Xodabande, I., Iravi, Y., Mansouri, B., & Matinparsa, H. (2022). Teaching academic words with digital flashcards: Investigating the effectiveness of mobile-assisted vocabulary learning for university students. *Frontiers in Psychology*, *13*, 893821. https://doi.org/10.3389/fpsyg.2022.893821