Indonesia Berdaya, 8(1) 2026: 63-82
https://doi.org/10.47679/ib.20261683

P-ISSN 2502-4825
E-ISSN 2502-9495

A Community-Based Approach to Improving English Writing Skills Through MALL and Morphological Instruction

Savira Zaniar[1], Tiyas Saputri[2], Edi Pujo Basuki[3], Mujad Didien Afandi[4], Elsa Fernanda[5], Sofiatul Nur Rachma Adi[6]

Published online: 24 November 2026

ABSTRACT

Many English learners can recognise words but struggle to build and use the correct word form in writing. This community engagement program, carried out with SMP Al Islah Surabaya under the Kemitraan Masyarakat scheme, combined Mobile-Assisted Language Learning (MALL) with explicit morphological instruction to strengthen the word-level foundations of 32 ninth-grade students' English writing. A diagnostic analysis of 615 morphological errors in the students' initial writing showed that derivational errors accounted for 76% and inflectional errors for 12%. The program followed a participatory cycle of needs analysis, pre-test, a face-to-face workshop, 7 weeks of WhatsApp-based practice with peer and teacher feedback, mentoring, post-test, and reflection. Using a one-group pre-test–post-test design, the program was evaluated with a morphological awareness test and a morphological mastery test, the latter measuring the use of correct word forms in sentences. Mean awareness scores increased from 59.84 to 75.31 and mastery scores from 61.25 to 76.25. Wilcoxon signed-rank tests showed that both gains were significant ($Z = -4.98$ and $Z = -4.90$, $p < .001$; $r = 0.88$ and 0.87), and the mean normalized gain was 0.40 on both tests, a medium gain on Hake's scale. No student's score declined, and the share of students in the high or very high band rose from 44% to 78% for awareness and from 47% to 84% for mastery. Because the design had no comparison group and overall writing quality was not scored, the results indicate within-group improvement in word-formation knowledge and use, one component of writing, rather than a causal effect on writing proficiency. The program offers a low-cost model for community-based English writing support that the partner teacher can sustain.

Keywords: English Language Acquisition; English Writing Skills; Mobile-Assisted Language Learning; Morphological Awareness; Morphology Instruction; Quality of Education (SDG 4)

INTRODUCTION

Writing is often the weakest of the four English skills among Indonesian junior high school students, and much of this weakness lies at the level of the word. Students may know the base word happy yet write "I feel happiness today" or "She sings happy." The idea is clear, but the word form does not fit the sentence. Such errors show that knowing a word's meaning is not the same as knowing how to change its form for different grammatical roles.

Morphological knowledge, the understanding of how words are built from roots and affixes, addresses this gap directly. A recent meta-analysis of 28 studies found that morphological instruction improved children's reading and spelling (Colenbrander et al., 2024), and studies with EFL learners report gains in vocabulary (Anggrisia & Billah, 2024) and in the written use of word forms (Khan, 2024; Wahid & Farooq, 2019). Knowing word families lets learners generate several related forms from one base (Nation, 2022). Yet in many schools and community learning settings, grammar is taught at the sentence level while word formation receives little systematic attention.

[1*-6] English Education Department, Faculty of Teacher Training and Education, Universitas Nahdlatul Ulama Surabaya, Indonesia

*) *corresponding author*

Savira Zaniar
Email: savira@unusa.ac.id

A second gap is limited practice time. Classroom hours are few, and learners rarely write outside class. Mobile-Assisted Language Learning (MALL) offers a way to extend practice, since most learners already own smartphones. Mobile devices support flexible, interactive learning beyond the classroom (Kukulska-Hulme & Viberg, 2018), and a recent meta-analysis of 65 studies found that mobile applications benefit vocabulary learning (Zhou & Zhou, 2026). Mobile tasks can deliver short, frequent morphology exercises and collect writing samples for feedback.

Although both lines of research are well established, they have seldom been brought together. Most morphology-instruction studies involve children in English-speaking countries (Colenbrander et al., 2024; Goodwin & Ahn, 2013) or EFL learners at senior high school and university level (Anggrisia & Billah, 2024; Asaad, 2024; Wahid & Farooq, 2019), and they deliver instruction face to face. MALL research, in turn, has focused on vocabulary in general, mostly measured as knowledge of word meaning, rather than on how learners build and use word forms (Zhou & Zhou, 2026). Few studies have examined whether mobile practice can extend explicit morphology instruction for junior high school EFL writers, and fewer still have done so in a community engagement setting, where a program must fit the partner school's existing resources and be sustained by its teachers after the university team leaves.

This program addresses that gap in three ways. First, the content of instruction was derived from a diagnostic analysis of the partner students' own writing errors rather than from a general morphology syllabus, so it concentrated on derivational morphology, the dominant source of their errors. Second, MALL was delivered through WhatsApp groups that students and teachers already used every day, so extended practice and feedback required no new devices, applications, or training. Third, outcomes were measured at two levels, morphological awareness (analysing word structure) and morphological mastery (producing the correct form in a sentence), so that the study could examine whether recognising word parts and using them in writing develop together. The article therefore contributes (a) evidence on MALL-supported morphology instruction for Indonesian junior high school learners, (b) a replicable, low-cost program model for school–university partnerships, and (c) data on the relationship between awareness and mastery in a short intervention.

Needs analysis with SMP Al Islah Surabaya, a junior high school in Surabaya, East Java, confirmed both problems. An analysis of the 32 students' initial writing (the coding procedure is described in the Materials and Methods section) showed that derivation was the dominant problem: 76% of morphological errors involved derivational affixes and word-class changes (e.g., using the noun "happiness" where the adjective "happy" is needed, or the adjective "happy" where the adverb "happily" is needed), while 12% involved inflection (e.g., missing plural -s or past tense -ed). The remaining 12% fell into other categories. Teachers also reported few resources for structured writing practice. These findings shaped the program's focus on derivational morphology.

This program therefore aimed to build participants' understanding of English word formation, provide mobile-supported practice linking morphology to sentence- and paragraph-level writing, and improve the accuracy of word forms in their writing, as part of the vocabulary and language-use components of writing quality (Jacobs et al., 1981). Overall writing quality was not scored in this study; the outcomes reported here are word-level morphological knowledge and use. The program also aimed to equip English teachers/facilitators at the partner site to continue the approach after the program ended.

To evaluate the program's effectiveness, this article addresses two research questions:

1. Is there a significant difference between students' morphological awareness and morphological mastery scores before (pre-test) and after (post-test) MALL-supported morphological instruction?
2. To what extent does MALL-supported morphological instruction improve students' morphological awareness and mastery, as measured by the normalized gain (N-gain) from the pre-test to the post-test?

LITERATURE REVIEW

Second Language Acquisition and Attention to Form

Second language acquisition depends on both input and output. Krashen (1982) argued that learners acquire language mainly through comprehensible input slightly beyond their current level, while Swain (1985) showed that producing language, especially in writing, pushes learners to notice gaps between what they want to say and what they can say. Input and output alone, however, may not be enough for forms that carry little meaning or are easy to overlook. Instruction that draws learners' attention to such forms helps them notice what they would otherwise miss and can speed up their acquisition (Ellis, 2008; Lightbown & Spada, 2021). Word-formation patterns are a clear example: learners meet *happy*, *happiness*, and *happily* in input but may not register how each form fits a different sentence slot. This supports combining meaningful writing practice with explicit attention to form.

Morphological Awareness and Morphological Mastery

Morphological awareness is the ability to reflect on and manipulate the smallest meaningful units of words, such as roots and affixes (Carlisle, 1995). It supports vocabulary growth, spelling, and comprehension (Kuo & Anderson, 2006), and knowing word families allows learners to recognise and produce related forms from one base (Bauer & Nation, 1993). In this article, morphological mastery refers to the productive side of this knowledge: the ability to choose and produce the correct derived or inflected form for a given sentence context. The distinction matters because the two need not develop at the same pace. Wang and Zhang (2023) showed that awareness is distinct from morphological knowledge: among 226 Chinese secondary EFL learners, knowledge of morphemes affected vocabulary only through awareness, which suggests that learners must actively analyse words, not just know a list of affixes. Derivational affixes are more numerous than inflectional ones and change a word's class (Plag, 2018), which makes derivation particularly demanding for learners.

Morphology Instruction

Morphology instruction teaches learners how words are built and how affixes change meaning and word class (Plag, 2018). Reviews show that such instruction improves literacy outcomes, with stronger effects for less advanced learners (Bowers et al., 2010; Carlisle, 2010), and Goodwin and Ahn (2013) reported moderate effects on vocabulary and spelling across 30 studies. A more recent meta-analysis of 28 studies found small-to-moderate effects on children's reading and spelling, strongest for the words that were taught, with considerable variation linked to intervention content and dosage (Colenbrander et al., 2024). Nation (2022) recommends teaching common affixes and word-family relationships explicitly, since this multiplies the vocabulary learners can use, and Nagy and Townsend (2012) describe morphological analysis as a key tool for academic vocabulary. Studies with EFL learners report similar benefits. Three weeks of morphological awareness instruction improved Indonesian senior high school students' vocabulary compared with traditional teaching (Anggrisia & Billah, 2024), and explicit instruction on irregular inflections improved Saudi EFL learners' use of those forms in writing (Khan, 2024). Most of this evidence, however, comes from face-to-face instruction with older learners or L1 children, and outcomes are usually measured as reading, spelling, or vocabulary rather than as word use in writing.

Morphology and Writing

Writing is a complex process of planning, translating ideas into text, and reviewing (Flower & Hayes, 1981). For EFL learners, this process is limited by their linguistic resources, so accuracy in grammar and vocabulary strongly affects writing quality, and feedback and peer review help learners

notice and revise their errors (Hyland, 2019). Analytic writing rubrics reflect this by scoring vocabulary (including word form) and language use as separate components of writing quality (Jacobs et al., 1981). Morphological knowledge contributes directly to these components. Asaad (2024) found that morphological awareness was linked to L2 postgraduates' academic writing through their vocabulary knowledge, and Wahid and Farooq (2019) found that undergraduate EFL students who received explicit morphology instruction outperformed a control group on word-formation tasks.

Mobile-Assisted Language Learning

Mobile-Assisted Language Learning (MALL) uses mobile devices to support learning that is personal, flexible, and not tied to the classroom. In a review of studies from 2012 to 2016, Kukulska-Hulme and Viberg (2018) found that features such as immediate response and social interaction make mobile devices well suited to collaborative language learning. Mobile tools also suit short, frequent tasks: in a meta-analysis of 65 experimental and quasi-experimental studies, Zhou and Zhou (2026) found that mobile applications improved vocabulary learning, with a medium effect after correcting for publication bias, and that apps designed for learning outperformed general-purpose apps. For writing, a systematic review of 21 studies identified easier access, greater involvement, stronger motivation, control over pace, learner independence, and interaction as the main benefits of mobile learning (Khamraeva Shavkatovna et al., 2024).

Conceptual Framework

Taken together, these strands suggest a pathway that guided the program design. Explicit morphology instruction makes affixes and their functions noticeable (Ellis, 2008; Lightbown & Spada, 2021) and so builds morphological awareness. Mobile tasks provide the repeated practice and quick feedback that derivational morphology requires (Kukulska-Hulme & Viberg, 2018; Nation, 2022), and because these tasks ask learners to produce the forms in their own sentences, they push awareness towards morphological mastery (Swain, 1985). Mastery, in turn, corresponds to the accurate use of word forms that analytic rubrics score under vocabulary and language use (Jacobs et al., 1981), one component of writing quality. The present study measures the first two links in this pathway, awareness and mastery. The link to writing is illustrated with examples from students' writing but was not measured with a scored writing test.

MATERIALS AND METHODS

Research Design

This community engagement program was evaluated with a quantitative, one-group pre-test–post-test design (Fraenkel et al., 2012), supplemented by a Likert-scale questionnaire describing students' responses to the program. The same group of 32 students was tested before the program (O_1), received MALL-supported morphological instruction (X), and was tested again after the program (O_2). The design has no control group. Without one, the change from O_1 to O_2 cannot be attributed to the program alone: it may also reflect regular English lessons, maturation, or familiarity with the test, and very low or very high initial scores tend to move towards the mean on retesting (Shadish et al., 2002). The design was chosen because the program was a community engagement activity for a whole class at the partner school, and a comparable class that did not receive the program was not available due to approval and code ethics from the partnership. Two steps reduce, but do not remove, these threats: the normalized gain accounts for each student's starting level, and the Wilcoxon signed-rank test with effect size checks that the change is unlikely to be due to chance. The results are therefore

interpreted as evidence of improvement within this group, not as proof of a causal effect (Creswell, 2012; Shadish et al., 2002).

Participants

The participants were all 32 ninth-grade students in one class at SMP Al Islah Surabaya. The class was selected purposively by the school partnership as the class with the greatest need in writing, and every student in it took part; no student was excluded from the analysis. The class consisted of 21 girls and 11 boys aged 14-16 years, who had studied English for about 8-10 years and had English lessons of 160 minutes per week. All students had access to a smartphone with WhatsApp, either their own or a parent's. Their pre-test scores ranged from 30 to 80, so the class covered all five score bands (Table 5). The school's English teacher took part in every stage as a partner. Participation in the tests and questionnaire was voluntary.

Program Procedure

The program followed the participatory cycle summarised in Table 3.1. It ran from June, 22nd 2026 to Nov, 3rd 2026 with 3 offline meetings in 19 weeks.

Table 1. Stages, timing, and activities of the program

Stage	Timing	Main activities	Participants
1.Needs analysis	online meeting / week	Diagnostic analysis of students' initial writing; discussion with the English teacher	Team, teacher, students
2. Pre-test	July, 10th 2026	Morphological awareness and mastery tests (O_1)	Students
3. Workshop	July, 10th 2026 July, 23rd 2026	Explicit instruction on roots, affixes, and word families; word-formation, sentence, and guided paragraph tasks	Team, teacher, students
4. Mobile practice	7 weeks; 2 tasks per 3 weeks	WhatsApp tasks with peer feedback and teacher/team feedback	Students, teacher, team
5. Mentoring	2 online sessions	Coaching the teacher to set and give feedback on tasks with the students	Team, teacher
6. Post-test	Oct, 2nd 2026	Same tests under the same conditions (O_2); questionnaire	Students
7. Reflection	Nov, 3rd 2026	Review of results with the school; module handed to the teacher	Team, teacher, school

Face-to-face workshop. The workshop lasted 60 minutes in every meeting and was led by the university team together with the English teacher. It introduced roots, prefixes, and suffixes and the word classes that common suffixes signal (e.g., -ness forming nouns and -ly forming adverbs), with most of the time given to derivation because of the needs analysis. Each affix was followed straight away by word-formation practice (building a word family from one base), a sentence task requiring the right form, and guided paragraph writing in which students checked the word forms in their own drafts.

Mobile practice through WhatsApp. After the workshop, the class continued in a WhatsApp group created for the program, with the teacher and team members as administrators for one class group. Two tasks were shared with the participants every three weeks, allowing them to have sufficient time to consider and reflect on vocabulary, lexical items, and grammatical structures in their writing carefully. The first task focused on their perspectives on word classes, with the aim of enhancing their understanding of morphology. In the second task, participants were asked to share their personal experiences related to daily routines or habitual activities. Samples of their writing were collected during the workshop and submitted as part of the final post-test.

Feedback procedure. Feedback was given in two layers. Students first commented on each other's answers in the peer-correction tasks. The team and teacher then gave corrective feedback within 3 weeks of each task: errors made by many students were explained in a summary message to the whole

group, and individual errors were addressed in a reply that pointed to the affix and asked the student to correct the form before the answer was given.

Mentoring and reflection. Mentoring consisted of 2 online sessions for 40-60 minutes of each. The sessions involved the university team and the school's English teacher and focused on three areas to design short word-formation and sentence tasks for the WhatsApp group, giving corrective feedback that directs students to the affix rather than supplying the correct form and using students' responses to decide which affixes required to be taught for the next session in class. In the later sessions, the teacher posted the tasks and gave feedback independently while the team observed and commented, so that responsibility for the mobile practice moved gradually from the team to the teacher. In the reflection stage, the team, the teacher, and the school reviewed the results, and the team handed over a morphology and writing module so that the teacher could continue the approach. In the reflection stage, the team, the teacher, and the school reviewed the results, and the team handed over a morphology and writing module so that the teacher could continue the approach.

Diagnostic Error Analysis

To describe the types of morphological errors students made, all incorrect responses on the pre-test and post-test were analysed. With 32 students, 30 items, and two tests, there were 1,920 responses in total, of which 615 were incorrect (381 at pre-test and 234 at post-test). An error was defined as a response in which the form of a word (its affix, word class, or inflection) did not fit the item or sentence. Each incorrect response was coded into one of three categories, such as **derivational errors**, involving a derivational affix or a change of word class, **inflectional errors**, involving inflectional suffixes such as plural -s, third-person singular -s, and past tense -ed and **other errors**, covering responses that fitted neither category, such as misspelt affixes, incorrect compounds, and blank or unclassifiable answers. However, other errors were not included in the analysis because this study focuses on inflectional and derivational morphemes. The categorisation of other errors will be described in more detail in a separate, more comprehensive research article. The responses were coded by [the first author] and checked by a second member of the team, and disagreements were resolved by discussion. The percentage for each category was calculated as the number of errors in that category divided by the total number of errors (N = 615), multiplied by 100: derivational 467/615 = 75.9%, inflectional 74/615 = 12.0%, and other 74/615 = 12.0% (Table 4.1). Because the analysis combines both tests, it describes the overall error profile of the class rather than the change in errors from pre-test to post-test.

Instruments

Data were collected with two tests and a Likert-scale questionnaire (Table 1). The same two tests were used at pre-test and post-test. The morphological awareness test (20 items) measured students' knowledge of inflectional and derivational morphemes. Part I contained 10 multiple-choice items on four indicators: recognising inflectional morphemes (items 1–2), recognising derivational morphemes (items 3–4), distinguishing inflectional from derivational morphemes (items 5–6; e.g., *colorful* vs. *colors*), and inferring word meaning through morphological analysis (items 7–10; e.g., *unbreakable*, *irreplaceable*). Part II contained 10 short-answer items on two indicators: segmenting words into roots and affixes and explaining the function of each affix (items 11–15; e.g., *reconstruction*, *carelessness*), and changing a base word into other word classes through affixation (items 16–20; e.g., *beauty* → *beautiful* → *beautifully*). The morphological mastery test (10 items) measured students' ability to produce and correct word forms in sentence context. Part I contained four sentence-completion items on inflection (e.g., "Every morning, she ___ (walk) to school"), Part II three word-formation items on derivation (e.g., "Her ___ (happy) was obvious when she received the award"), and Part III three error-correction items in which students found and corrected a morphological error (e.g., "My brother is a successed businessman").

Table 2 Instruments, scoring, and reliability

Instrument	Items and format	What it measures	Scoring	Reliability
Morphological awareness test	20 items: Part I, 10 multiple-choice; Part II, 10 short answers	Recognising and distinguishing inflectional and derivational morphemes; inferring meaning from morphemes; segmenting words; forming related words by affixation	Correct/incorrect; 5 points per item; maximum 100	$\alpha = .78$
Morphological mastery test	10 items: 4 sentence completion (inflection), 3 word formation (derivation), 3 error correction	Producing and correcting word forms in sentence context	Correct/incorrect; 10 points per item; maximum 100	$\alpha = .58$
Questionnaire	12 Likert items (4-point: 1 = strongly disagree to 4 = strongly agree)	Perceived usefulness of the morphology instruction, WhatsApp tasks, and feedback; overall confidence	Number and percentage of students choosing each response	Not estimable (ceiling effect); descriptive use only

Each item was scored as correct or incorrect. For short-answer, sentence-completion, word-formation, and error-correction items, an answer was accepted as correct when it matched the answer key or an acceptable alternative listed in the scoring guide, because students' answers could differ in form or spelling while still being morphologically correct; an error-correction item was accepted only when the error was identified and the correct form was written. Each awareness item was worth 5 points and each mastery item 10 points, giving a maximum score of 100 on each test. We estimated internal consistency from all 32 participants' pre-test responses using Cronbach's alpha, which is equivalent to KR-20 for dichotomously scored items. The coefficients were .78 for the awareness test and .58 for the mastery test. Coefficients of .70 or higher are generally regarded as adequate for research purposes (Fraenkel et al., 2012). The awareness test met this criterion. The mastery coefficient was lower, which is partly expected for a test of only 10 items, since alpha rises with test length; by the Spearman–Brown formula, a 20-item version with items of similar quality would reach about .73. One mastery item (item 10, error correction) correlated negatively with the rest of the test ($r = -.10$) and should be revised in future use. Mastery scores are therefore interpreted with more caution than awareness scores.

Neither test scores overall writing quality. The mastery test measures accurate word-form use in sentence context, which corresponds to the word-form element of the vocabulary and language use components of analytic writing rubrics (Jacobs et al., 1981). Claims about writing in this article are therefore limited to this component.

Both tests were scored on a 0–100 scale and grouped into five bands: very low (below 40), low (40–59), medium (60–69), high (70–79), and very high (80–100). A 12-item, four-point Likert-scale questionnaire (1 = strongly disagree, 2 = disagree, 3 = agree, 4 = strongly agree) was used to describe students' responses to the program. It covered four areas: the morphology instruction (items 1–4), mobile practice through WhatsApp (items 5–8), feedback (items 9–10), and overall perceptions (items 11–12). The scale had no neutral midpoint, and responses are reported as the number and percentage of students choosing each option. Internal consistency could not be meaningfully estimated for the questionnaire: responses clustered at the top of the scale (item means 3.44–3.94; total scores 42–48 of a possible 48, SD = 1.44), leaving too little variation between students, and Cronbach's alpha was negative. The questionnaire is therefore used only to describe students' perceptions, not as a measure in any statistical analysis.

Data Collection Procedures

Data were collected in three steps: the pre-test before the workshop, the post-test after the mobile practice period, and the questionnaire after the post-test. The pre-test and post-test used the same items under the same conditions (45 minutes, in class, supervised by the teacher and the team), so that changes in scores did not reflect a different instrument. The interval between the two tests was 12 weeks. The questionnaire was completed after the post-test so that responses could refer to the whole program.

Data Analysis

Quantitative data were analysed with descriptive and inferential statistics. Descriptive statistics (means, standard deviations, minimum and maximum scores, frequencies, and percentages) summarised students' levels of morphological awareness and mastery at pre-test and post-test and their distribution across the score bands. Inferential statistics were used to test whether the pre-test–post-test differences were significant and to examine the relationship between awareness and mastery. Questionnaire responses were reported as percentages of students who agreed with each statement.

The effectiveness of the program was determined with the normalized gain (N-gain) score (Hake, 1998). N-gain compares the improvement a student actually made with the maximum improvement that student could have made, so a student who starts high is not penalised for having less room to grow. It is calculated as:

$$\text{N-Gain} = \frac{\text{Post} - \text{Pre}}{\text{Maks} - \text{Pre}}$$

Picture 1. N-Gain Score Formulation

N-gain was calculated for each student, and the mean N-gain of the 32 students was used to judge the program's effectiveness for each test. For comparison with Hake's (1998) original procedure, the class-average gain was also calculated from the pre-test and post-test means. N-gain values were interpreted with Hake's (1998) categories (Table 3). These categories were used as the only interpretive scale because they are defined in the source that introduced the normalized gain. Percentage-based effectiveness bands (for example, below 40% as "not effective"), which are common in Indonesian studies, were not used, because their thresholds could not be traced to a primary source and because they classify the same gain differently from Hake's scale.

Table 3 N-gain interpretation (Hake, 1998)

N-gain (g)	Category
$g \geq 0.70$	High
$0.30 \leq g < 0.70$	Medium
$g < 0.30$	Low

N-gain shows how large the improvement was but not whether it could have occurred by chance. The significance of the pre-test–post-test difference was therefore checked with an inferential test. The Shapiro–Wilk test showed that the score differences were not normally distributed (awareness $W = 0.88$, $p = .002$; mastery $W = 0.84$, $p < .001$), so the Wilcoxon signed-rank test was used, with effect size $r = Z/\sqrt{N}$ (Fritz et al., 2012), interpreted with Cohen's (1988) thresholds of 0.10 (small), 0.30 (medium), and 0.50 (large). Pearson correlation was used to examine the relationship between awareness and mastery scores.

RESULTS AND DISCUSSION

Overall, the program produced significant improvements in both morphological awareness and morphological mastery, with large effect sizes and a mean N-gain of 0.40 for both tests, in Hake's (1998) medium category. Because these are within-group gains on word-level measures, they are interpreted below with the limits of the design in mind. The results are presented in four parts: the error profile of students' test responses, the answers to the two research questions, examples from students' writing, and students' responses to the questionnaire. The Discussion then interprets these findings in relation to earlier studies.

Morphological Errors in Students' Test Responses

Of the 1,920 responses on the pre-test and post-test, 615 (32.0%) were incorrect: 381 of 960 at pre-test (39.7%) and 234 of 960 at post-test (24.4%), a mean of 19.2 errors per student across the two tests. As Table 4 shows, derivational errors accounted for 467 of these errors (75.9%) and inflectional errors for 74 (12.0%). The remaining 74 errors (12.0%) fell outside these two categories and, as explained in the Materials and Methods section, were not analysed further. The most common derivational errors involved choosing the wrong member of a word family, such as a noun where an adjective was needed (*happiness* for *happy*) or an adjective where an adverb was needed (*happy* for *happily*). Students usually knew the base word; what they had not yet learned was which form the sentence required. Inflectional errors were far fewer and mostly involved plural -s, third-person singular -s, and past tense -ed. The number of incorrect responses also fell from 381 at pre-test to 234 at post-test, a reduction of 38.6%, which is consistent with the score gains reported under research question 1. The share of derivational errors should be read in light of the test's composition: 19 of the 30 items targeted derivation, 8 targeted inflection, and 3 covered both. Even allowing for this, derivation remained harder: relative to the number of single-category items, there were about 0.38 derivational errors per derivation-item response (467 ÷ [19 × 64]) compared with about 0.14 inflectional errors per inflection-item response (74 ÷ [8 × 64]). These rates are approximate, because errors on the three mixed items are included in the counts.

Table 4 Morphological errors in students' pre-test and post-test responses (N = 615)

Error category	Description and example	n	%
Derivational	Wrong derivational affix or word class, e.g., happiness for happy (noun for adjective); happy for happily (adjective for adverb)	467	75.9
Inflectional	Missing or wrong inflectional suffix, mainly plural -s, third-person singular -s, and past tense -ed	74	12.0
Other	Errors outside the two categories; not analysed further in this study	74	12.0
Total		615	100.0

Note. 32 students × 30 items × 2 tests = 1,920 responses; 381 errors at pre-test and 234 at post-test.

This pattern suggests that students had largely acquired basic inflection but had not yet learned how affixes change a word's class. Derivational affixes are more numerous than inflectional ones and change a word's class (Plag, 2018), which makes them harder to acquire from exposure alone. The dominance of derivational errors therefore supports the program's decision to give derivation the largest share of workshop time and mobile tasks.

Research Question 1: Difference between Pre-test and Post-test Scores

The first research question asked whether there was a significant difference between students' morphological awareness and mastery scores before and after MALL-supported morphological instruction. Table 5 answers this question directly. Tables 4.3 to 4.6 then show how the difference

was distributed across the class, how individual students moved between score bands, how gains varied with students' starting level, and how the two tests related to each other.

Table 5. Descriptive statistics and Wilcoxon signed-rank test results for pre-test and post-test scores (N = 32)

Measure	Pre-test M (SD)	Post-test M (SD)	Median pre → post	Min–max pre → post	Mean gain	Wilcoxon Z	r
Morphological awareness	59.84 (15.68)	75.31 (12.11)	60.0 → 77.5	30–80 → 45–100	15.47	−4.98*	0.88
Morphological mastery	61.25 (14.97)	76.25 (12.38)	60.0 → 80.0	30–80 → 50–100	15.00	−4.90*	0.87

*p < .001

Table 5 shows that the difference between pre-test and post-test scores was significant on both tests. Mean awareness scores increased by 15.47 points and mean mastery scores by 15.00 points, and the medians increased by 17.5 and 20.0 points. The Wilcoxon signed-rank test, used because the score differences were not normally distributed, gave $Z = -4.98$ for awareness and $Z = -4.90$ for mastery, both with $p < .001$. A p-value this small means that, if the instruction had made no difference, a change of this size would occur by chance less than once in 1,000 cases. The answer to research question 1 is therefore yes: students scored significantly higher after the instruction than before it.

The effect sizes show that the difference also matters in practice. Values of $r = 0.88$ and $r = 0.87$ are well above 0.50, the conventional threshold for a large effect (Cohen, 1988). The improvement was also consistent. All 32 students raised their awareness score. In mastery, 30 students improved and two kept the same score of 80, so no student declined on either test. The weakest students also improved: the lowest score increased from 30 to 45 in awareness and from 30 to 50 in mastery. Finally, scores became less spread out. Standard deviations fell from 15.68 to 12.11 in awareness and from 14.97 to 12.38 in mastery, showing that the gap between stronger and weaker students narrowed.

Table 6. Number of students in each score band (N = 32)

Score band	Awareness pre-test	Awareness post-test	Mastery pre-test	Mastery post-test
Very high (80–100)	3	16	6	20
High (70–79)	11	9	9	7
Medium (60–69)	7	4	8	2
Low (40–59)	8	3	7	3
Very low (below 40)	3	0	2	0

Table 6 shows what the significant difference looks like across the class. Before the instruction, 14 students (44%) in awareness and 15 (47%) in mastery were in the high or very high bands; afterwards, 25 (78%) and 27 (84%) were. The very high band grew most, from 3 to 16 students in awareness and from 6 to 20 in mastery, and the very low band was empty on both tests. At the lower end, the number of students below 60 fell from 11 to 3 in awareness and from 9 to 3 in mastery. The difference found in Table 4.2 therefore reflects a shift of the whole class, not large gains by a few students.

Table 7 follows each student from the pre-test band (rows) to the post-test band (columns); numbers on the diagonal are students who stayed in the same band. It shows that upward movement was the rule. In awareness, 28 of the 32 students moved up at least one band and 6 moved up two or more. The only students who stayed in the same band were the three already in the very high band and one student in the low band. In mastery, 25 students moved up, 9 of them by two or more bands, and the seven who stayed were the six already in the very high band and one student in the low band. No student on either test moved down.

Table 7. Movement of students between score bands from pre-test to post-test (N = 32)

Pre-test band (n)	Very low	Low	Medium	High	Very high
Morphological awareness: post-test band					
Very low (3)	–	2	–	1	–
Low (8)	–	1	4	3	–
Medium (7)	–	–	–	5	2
High (11)	–	–	–	–	11
Very high (3)	–	–	–	–	3
Morphological mastery: post-test band					
Very low (2)	–	2	–	–	–
Low (7)	–	1	2	4	–
Medium (8)	–	–	–	3	5
High (9)	–	–	–	–	9
Very high (6)	–	–	–	–	6

Every student who started in the high band finished in the very high band (11 in awareness and 9 in mastery), so the instruction helped strong students reach the top level, not only weaker students to catch up. All students who began in the very low band left it, but four of the five were still below 60 at post-test, which suggests that the weakest students need more time than one program provided.

Table 8. Mean gains by students' pre-test level (N = 32)

Pre-test level	Awareness: n	Awareness mean gain (points)	Awareness mean N-gain	Mastery: n	Mastery mean gain (points)	Mastery mean N-gain
Below 60	11	20.91	0.36	9	20.00	0.34
60–79	18	12.78	0.40	17	14.71	0.42
80 and above	3	11.67	0.58	6	8.33	0.42
All students	**32**	**15.47**	**0.40**	**32**	**15.00**	**0.40**

Table 8 groups students by pre-test score. It shows that point gains were largest for the students who started lowest. Students who scored below 60 at pre-test gained an average of 20.91 points in awareness and 20.00 in mastery, compared with 11.67 and 8.33 points for those who started at 80 or above. Pre-test scores were negatively correlated with point gains ($r = -0.67$ for awareness and $r = -0.57$ for mastery, $p \leq 0.001$). Part of this pattern reflects the room each group had to improve: a student who starts at 80 can gain at most 20 points, while a student who starts at 40 can gain 60. The N-gain columns, which take this room into account, are discussed under research question 2.

Table 9. Comparison and correlation of morphological awareness and mastery scores (N = 32)

Comparison	Pre-test	Post-test	Gain
Mean awareness score	59.84	75.31	15.47
Mean mastery score	61.25	76.25	15.00
Difference between tests (Wilcoxon p)	0.46	0.24	0.97
Correlation between tests (Pearson r)	0.85*	0.93*	0.42**

*p < .001. **p = .017

Table 9 compares the two tests. Students' awareness and mastery scores did not differ significantly at pre-test ($p = .46$) or at post-test ($p = .24$), and the gains on the two tests were almost identical (15.47 and 15.00 points, $p = .97$). Mastery scores were in fact slightly higher than awareness scores on both occasions. The two tests were strongly correlated, with $r = .85$ at pre-test and $r = .93$ at post-test ($p < .001$), so students who better understood word structure also used word forms more accurately. Students' gains on the two tests were moderately correlated as well ($r = .42$, $p = .017$): those who improved more in awareness also tended to improve more in mastery.

In answer to research question 1, there was a significant difference between students' pre-test and post-test scores on both tests. The difference was large (r = 0.88 and 0.87), consistent across students, spread across the whole class, and similar on both tests.

Research Question 2: Extent of Improvement

The second research question asked how much the instruction improved students' morphological awareness and mastery. The normalized gain answers this by comparing each student's gain with the room that student had to improve. Table 4.7 presents the results on Hake's (1998) scale.

Table 10. Distribution of students' individual N-gain scores by Hake's (1998) categories (N = 32)

N-gain category	Awareness: students (%)	Mastery: students (%)
High ($g \geq 0.70$)	1 (3%)	2 (6%)
Medium ($0.30 \leq g < 0.70$)	23 (72%)	22 (69%)
Low ($g < 0.30$)	8 (25%)	8 (25%)
Mean N-gain (SD)	**0.40 (0.16), medium**	**0.40 (0.20), medium**

Table 10 shows that the improvement was moderate. The mean N-gain was 0.40 on both tests, in Hake's (1998) medium category. An N-gain of 0.40 means that, on average, students closed 40% of the gap between their pre-test score and the maximum score of 100. Most students fell in the medium category: 23 (72%) in awareness and 22 (69%) in mastery. One student in awareness and two in mastery reached a high gain, while a quarter of the class on each test made a low gain. The class-average gain calculated from the group means, as in Hake's (1998) original procedure, was 0.39 for awareness ((75.31 − 59.84)/(100 − 59.84)) and 0.39 for mastery ((76.25 − 61.25)/(100 − 61.25)), also in the medium category, so the classification does not depend on how the gain is averaged.

The low-gain group was not limited to one ability level. The eight students with low gains on each test had pre-test scores ranging from 30 to 80. N-gain was also not significantly related to pre-test scores (r = .23, p = .20 for awareness; r = .16, p = .38 for mastery), and the mean N-gain in Table 4.5 was similar across the three starting levels, from 0.34 to 0.42 in mastery and from 0.36 to 0.40 in awareness for students below 80. (The higher awareness value of 0.58 for students at 80 or above is based on only three students.) Once each student's room for improvement is taken into account, weaker and stronger students therefore benefited to a similar degree. This differs from the point gains reported under research question 1, which favoured weaker students mainly because they had more room to grow.

Table 10 and the class-average gain give a consistent answer to research question 2: the instruction produced a medium normalized gain, with students closing on average about two fifths of the distance between their pre-test score and the maximum. The gains were real and widespread, as research question 1 shows, but most students still had considerable room to improve after a short program, so the extent of improvement is best described as moderate.

Evidence from Students' Writing

To show how the gains appeared in students' own writing, Table 11 compares sentences written by three students during the workshop with sentences on similar content from their second WhatsApp task, in which they wrote about their daily routines. The students were chosen to represent low, medium, and high pre-test levels.

The three examples show different kinds of change. Before the program, the low-level student (S03) left the verb uninflected in a past-time narrative (*The party start*); afterwards, the student marked the past tense with -ed (*the party started*). The medium-level student (S16) placed the participial adjective *excited* at the end of the clause, where an adverb (*excitedly*) would be needed; afterwards, the student used the -ing adjective correctly after a linking verb (*Lunchtime is exciting*), choosing the form that describes the event rather than the person's feeling. The high-level student

(S18) made no morphological error before the program but used a wider range of derived and inflected forms afterwards, including the adverb *finally* (final + -ly) and the participle *dressed*; the change from *get out from* to *get out of* in the same sentence is a prepositional rather than a morphological correction. These examples illustrate the kind of change measured by the mastery test. They were not scored, however, and do not show change in overall writing quality.

Table 11. Examples of students' sentences before and after the program

Student (pre-test level)	Before the program	After the program	Change in word form
S03 (low)	The party start.	Then, the party started.	Inflection: missing past tense → -ed added (start → started)
S16 (medium)	I enjoy my lunch with friends in the canteen excited.	Lunchtime is exciting for me. I enjoy my lunch with friends in the canteen, this is my school life!	Derivation: misplaced -ed adjective (excited) → -ing adjective after a linking verb (exciting)
S18 (high)	When I get out from my bed, I go to the bathroom.	When I finally get out of bed, I go to the washroom, wash up, get dressed, and brush my hair.	Wider range of forms: derived adverb finally (-ly), participle dressed (-ed)

Students' Responses to the Program

Table 12. Students' responses to the questionnaire (N = 32)

Statement	Strongly agree, n (%)	Agree, n (%)	Disagree, n (%)	Strongly disagree, n (%)
Morphology instruction				
1. The explanation of roots, prefixes, and suffixes helped me understand how English words are formed.	24 (75.0)	8 (25.0)	0 (0.0)	0 (0.0)
2. I can now tell whether a word is a noun, adjective, verb, or adverb by looking at its suffix.	18 (56.2)	10 (31.2)	4 (12.5)	0 (0.0)
3. Building word families (e.g., happy – happiness – happily) helped me remember new words.	20 (62.5)	10 (31.2)	2 (6.2)	0 (0.0)
4. I check the word form more carefully when I write sentences in English.	28 (87.5)	4 (12.5)	0 (0.0)	0 (0.0)
Mobile practice through WhatsApp				
5. The WhatsApp tasks helped me practise English outside the classroom.	20 (62.5)	10 (31.2)	2 (6.2)	0 (0.0)
6. Three weeks gave me enough time to think about the word forms in my writing.	15 (46.9)	16 (50.0)	1 (3.1)	0 (0.0)
7. Writing about word classes in the WhatsApp group made the lesson easier to understand.	27 (84.4)	5 (15.6)	0 (0.0)	0 (0.0)
8. Writing about my daily routine helped me use new word forms in real sentences.	25 (78.1)	7 (21.9)	0 (0.0)	0 (0.0)
Feedback				
9. Comments from my classmates helped me notice my mistakes.	27 (84.4)	4 (12.5)	1 (3.1)	0 (0.0)
10. Feedback from the teacher and team helped me correct the word forms in my writing.	30 (93.8)	2 (6.2)	0 (0.0)	0 (0.0)
Overall perceptions				
11. I feel more confident writing English sentences after this program.	25 (78.1)	7 (21.9)	0 (0.0)	0 (0.0)
12. I would like to continue learning English with WhatsApp tasks like these.	23 (71.9)	9 (28.1)	0 (0.0)	0 (0.0)

Note. 4-point scale: 1 = strongly disagree, 2 = disagree, 3 = agree, 4 = strongly agree.

Table 12 summarises students' responses to the 12-item Likert-scale questionnaire (N = 32), grouped into four areas: morphology instruction, mobile practice through WhatsApp, feedback, and

overall perceptions. Responses were strongly positive: no student chose strongly disagree for any statement, and agreement (agree or strongly agree) ranged from 87.5% to 100%. Agreement was highest for overall perceptions (100% on both items, including feeling more confident writing English sentences) and feedback (98.4% of responses), followed by mobile practice through WhatsApp (97.7%) and morphology instruction (95.3%). The statement with the most strong agreement was that feedback from the teacher and team helped students correct word forms (item 10; 30 students, 93.8%). Agreement was lowest for telling a word's class from its suffix (item 2; 87.5%, with four students disagreeing), which suggests that this skill remained difficult for some students. Fewer than half of the students strongly agreed that three weeks gave them enough time for each task (item 6; 15 students, 46.9%), although all but one agreed. These responses are self-reports and are presented as context for the test results, not as evidence of learning.

DISCUSSION

Significant improvement in awareness and mastery

The significant pre-test–post-test differences agree with evidence that explicit morphological instruction improves learners' knowledge and use of words (Anggrisia & Billah, 2024; Colenbrander et al., 2024), but the size of the effects needs careful reading. The effect sizes here (r = 0.88 and 0.87) are much larger than the small-to-moderate effects reported in meta-analyses (Colenbrander et al., 2024; Goodwin & Ahn, 2013). The difference lies largely in design. Meta-analytic effects compare treated groups with control groups and so exclude gains from regular lessons, maturation, and test practice; a within-group effect includes all of these. Our effect sizes therefore show how consistently students' scores rose, not how much the program added beyond what students would have gained anyway, and they should not be compared directly with controlled effects. Our evidence is also narrower than that of Anggrisia and Billah (2024), who compared morphology instruction with traditional teaching and measured vocabulary, and of Wahid and Farooq (2019), whose undergraduate participants were compared with a control group. The present study extends this work to younger learners and to word-form use in sentences, but without the comparison that would support a causal claim.

The focus on derivation also distinguishes this program from earlier EFL studies. Khan (2024) targeted irregular inflection and found gains in learners' written use of those forms. In our partners' test responses, by contrast, inflection accounted for only 12% of morphological errors, which suggests that these learners had largely acquired the inflectional system and that derivation, with its many affixes and class changes, was the harder system (Plag, 2018). Focusing on derivation, the main source of the partners' difficulties, rather than following a general syllabus is therefore one reason the instruction could produce consistent gains in a short time.

The findings also fit the view that instruction which draws learners' attention to form helps with features they rarely notice on their own (Lightbown & Spada, 2021). Before the program, students met words such as *happy*, *happiness*, and *happily* in their reading, yet still chose the wrong form when writing. Making each affix and its function visible, and then asking students to use the form in their own sentences, turned this exposure into knowledge they could apply. Peer correction in the WhatsApp group added a further step: students had to notice an error in a classmate's writing and explain the correct form, which required the same kind of analysis that the awareness test measured. Noticing was not measured directly, however, so this mechanism is an interpretation consistent with the results rather than a finding.

The narrower spread of scores and the larger point gains among weaker students suggest that the instruction was especially useful for students who started with little explicit knowledge of word formation. For them, patterns such as -ness forming nouns and -ly forming adverbs were a new tool, whereas stronger students may already have picked up some of these patterns implicitly. This

interpretation should be read with care, however, because students with very low pre-test scores tend to score closer to the average on a second test, and students who already scored 80 had less room to improve. The similar N-gains across starting levels in Table 8 suggest that, relative to their room for growth, all groups benefited to a comparable degree.

The relationship between morphological awareness and mastery

Morphological awareness and mastery increased together. The two tests were strongly correlated at both pre-test and post-test, and students' gains on them were related. This fits Wang and Zhang's (2023) finding that knowledge of morphemes affects vocabulary through morphological awareness, and Asaad's (2024) finding that morphological awareness is linked to writing through vocabulary knowledge. In practical terms, the students who learned to see a word as a root plus affixes were also the students who chose the correct form in a sentence. Our data, however, are correlational and come from a single group, so they cannot show the direction of the relationship that Wang and Zhang (2023) tested with older secondary learners and a vocabulary outcome. Asaad's (2024) participants were postgraduates writing academic texts, whereas our students wrote sentences and short paragraphs. The very strong correlation at post-test ($r = .93$) may also partly reflect the shared content of the two tests, which targeted some of the same derivational suffixes (e.g., -ness and -ful), rather than a separate developmental link.

It is often assumed that recognising word parts develops well before the ability to use them. In this program, however, students did not score significantly lower on the mastery test than on the awareness test, either before or after the instruction (Table 9). One likely reason is the design of the instruction: each affix was linked to a sentence-level task from the start, through word-formation practice, guided paragraph writing, and the two WhatsApp writing tasks, so recognition and use were practised together rather than in sequence.

Moderate extent of improvement

The N-gain results show that the improvement, though significant, was moderate. The length of the program is the most likely reason. The instruction consisted of three 60-minute face-to-face meetings and seven weeks of mobile practice in which students completed two writing tasks, a short dose compared with many intervention studies. Colenbrander et al. (2024) found that effects varied widely with intervention content and dosage, and Zhou and Zhou (2026) reported larger effects for mobile vocabulary treatments lasting ten weeks or more. Derivational morphology is also demanding: affixes are numerous and each changes a word's class in its own way (Plag, 2018), so learners need repeated encounters to master them (Nation, 2022). Within a short program, students learned enough to raise their scores reliably but not enough to close most of the gap to full mastery.

A medium gain is in line with what Hake (1998) reported for interactive-engagement courses, whose class-average gains fell mostly in the medium range, whereas traditional courses produced low gains. The comparison is only indicative, since Hake's data come from physics courses and conceptual tests, but it suggests that a short, interactive program can reach a medium gain when it is tightly focused on learners' main errors. The same instruction delivered over more sessions, with more regular mobile practice, could be expected to produce larger normalized gains, and because N-gain did not depend on students' starting level, such an extension could benefit the whole class rather than one group of students.

The contribution of MALL

MALL contributed mainly by extending the amount and duration of practice. The two WhatsApp tasks carried the program beyond three face-to-face meetings into seven weeks of practice, and each task served a different purpose: the first asked students to reflect on word classes, strengthening their awareness of how affixes work, while the second asked them to write about their daily routines, which

required them to use the target forms in personal writing. The three-week response period for each task gave students time to check vocabulary and word forms before posting, and the group format allowed both peer comments and teacher feedback. This is consistent with the flexible, interactive nature of mobile learning (Kukulska-Hulme & Viberg, 2018) and with meta-analytic evidence that mobile applications aid vocabulary development (Zhou & Zhou, 2026). Because students already used WhatsApp every day, it required no new devices or training, and it made individual feedback practical for a group of 32. Students' questionnaire responses point the same way: 97.7% of responses to the WhatsApp items and 98.4% of responses to the feedback items were agreement. Because the responses were uniformly positive and self-reported, however, they show how students perceived the program rather than how much each component contributed to their gains.

These benefits match those identified in a systematic review of mobile learning for EFL writing, which found easier access, greater involvement, stronger motivation, control over pace, learner independence, and interaction to be the main advantages (Khamraeva Shavkatovna et al., 2024). At the same time, WhatsApp is a general-purpose messaging app, and Zhou and Zhou (2026) found that apps designed specifically for learning outperformed general-purpose apps. Future programs could therefore test whether a dedicated vocabulary or quiz application adds to the gains. With only two tasks over seven weeks, the amount of mobile practice was also modest, and more frequent, shorter tasks might strengthen its contribution. The design of this study, however, does not separate the effect of the mobile tasks from that of the workshop.

Implications for teachers and community engagement

The findings have three practical implications for English teachers at the partner school and in similar settings. First, derivational morphology deserves explicit attention in junior high school writing lessons, since it was the main source of word-form errors and responded well to direct teaching. Second, analysis and production should be taught together: identifying the suffix in *kindness* is most useful when it is followed straight away by writing a sentence that needs *kindness* rather than *kind*. Third, short mobile tasks on platforms students already use are a low-cost way to extend practice. For community engagement, the program also shows how a university and a school can work as partners: the teacher took part throughout and received a module to continue the approach, which supports quality education (SDG 4) by building local capacity rather than offering a one-off intervention.

Limitations

Several limitations should be considered when interpreting these results. First, the one-group pre-test–post-test design had no control group, so part of the gain may reflect regular English lessons, maturation, familiarity with the tests, or regression towards the mean, and the results cannot be read as a causal effect of the program. The interval of about 12 weeks between the pre-test (10 July 2026) and the post-test (2 October 2026) also included regular English lessons of 160 minutes per week, which may have contributed to the gains. Second, overall writing quality was not scored: the two tests measured word-level morphological knowledge and use, and the writing examples are illustrative, so the findings concern one component of writing rather than writing proficiency as a whole. Third, the design does not separate the contribution of the mobile tasks from that of the workshop. Fourth, the sample was a single class of 32 students at one school, which limits generalisation, and the questionnaire data are self-reports. In addition, the mastery test had lower internal consistency (α = .58) than the awareness test (α = .78), so the mastery results are less precise. The questionnaire also showed a strong ceiling effect, so its reliability could not be established and its results are descriptive only. Finally, there was no delayed post-test, and the analysis used total scores rather than item- or affix-level data, so it cannot show whether the gains were retained or which affixes students learned best. Future programs should address these points with a comparison class, a scored writing task

assessed with an analytic rubric before and after the program, a delayed post-test, and item-level analysis.

CONCLUSION

This community engagement program combined MALL with explicit, derivation-focused morphological instruction for 32 ninth-grade students at SMP Al Islah Surabaya. Students' morphological awareness rose from 59.84 to 75.31 and their morphological mastery from 61.25 to 76.25. Both gains were significant, no student declined, and the mean normalized gain of 0.40 on both tests was medium on Hake's scale. Awareness and mastery were strongly correlated and improved together, and gains relative to each student's room for growth were similar across starting levels.

The study contributes a program model in which the content of instruction comes from a diagnostic analysis of the partner students' own errors and practice is extended through a mobile platform already in daily use, and it adds evidence on MALL-supported morphology instruction for junior high school EFL learners, a group seldom studied. For teachers, the practical implications are to teach derivational morphology explicitly, to pair the analysis of word parts with their immediate use in sentences, and to use short mobile tasks to extend practice. The morphology and writing module handed to the school's English teacher allows the approach to continue after the program.

These conclusions are limited by the one-group design, the single class, and the absence of a scored measure of overall writing quality. The results show within-group improvement in word-formation knowledge and use, one component of writing, rather than a proven effect on writing proficiency. Future programs should run longer and include a comparison class, a delayed post-test, and an analytic writing assessment, to test whether these gains carry over into the quality of students' writing.

Acknowledgments

Our team shows our thanks and gratitude to the Directorate of Research, Innovation and Community Service for funding this program under the Kemitraan Masyarakat scheme, Universitas Nahdlatul Ulama Surabaya (UNUSA) for institutional support, and SMP Al Islah Surabaya, its principal M. Nasrulloh, its English teacher, and its Grade IX students for their participation.

DECLARATIONS

Ethics Approval and Consent to Participate

The program was reviewed and approved by the Directorate Research, Innovation and Community Service of Universitas Nahdlatul Ulama Surabaya. Institutional permission to conduct the program and collect data was granted by the principal of SMP Al Islah Surabaya. Because the participants were junior high school students under 18 years of age, written informed consent was obtained from each student's parent or legal guardian, and each student gave assent before taking part. Students and parents were told that participation was voluntary, that students could withdraw at any time without any effect on their school grades, and that test scores, questionnaire responses, and writing samples would be used only in anonymised form. All data were coded (e.g., S1, S2), and no names or other identifying details of students appear in this article.

Consent for Publication

Not applicable. The article contains no identifiable personal data of the student participants; writing examples are reported under anonymous codes.

Availability of Data and Materials

The anonymised dataset (pre-test, post-test, and N-gain scores for each student), the morphological awareness and mastery tests, the answer key and scoring guide, and the questionnaire are available from the corresponding author upon reasonable request. Students' original writing samples are not publicly available because they could identify participants who are minors. The morphology and writing module produced by the program can be obtained from the corresponding author.

Conflicts of Interest

The authors have no conflicts of interest to declare.

Funding

This work was funded by the Directorate of Research, Innovation and Community Service Universitas Nahdlatul Ulama Surabaya under the Kemitraan Masyarakat (Community Partnership) scheme in 2026.

Artificial Intelligence-Assisted Technology

During the revision of this manuscript, the authors used Claude (Anthropic) to help restructure sections in response to the reviewers' comments, improve the clarity of the English, and check that in-text citations matched the reference list. The tool was not used to generate, analyse, or interpret data. The authors reviewed and edited all AI-assisted text, verified every reference against publisher or indexing records, and take full responsibility for the content of the article.

Authors' Contributions

Team (Name)	Position	Field of Expertise	Task Description
Savira Zaniar, S.Hum., M.Hum.	Leader	English Education and Linguistics	Coordinating the program, supervising implementation, conducting morphology training, and preparing publications
Edi Pujo Basuki, S.Pd., M.Pd & Dr. Tiyas Saputri, S.S., M.Pd.	Member	English Education (English Teaching Writing and Assessment)	Designing instructional materials and assisting evaluation activities
Mujad Didien Afandi, S.S., M.Pd.	Member	Educational Technology and Second Language Acquisition	Developing MALL-based learning media and digital integration
Student Assistants (Elsa Fernanda and Sofiatul Nur Rachma Adi)	Supporting Team & Administrator	English Education	Documentation, data collection, technical support, and administration

All authors read and approved the final manuscript.

Authors' Information

Savira Zaniar, Tiyas Saputri, Edi Pujo Basuki and Mujad Didien Afandi are the lecturers in the English Education Department, Faculty of Teacher Training and Education, Universitas Nahdlatul Ulama Surabaya, Indonesia. On the other hand, **Elsa Fernanda and Sofiatul Nur Rachma Adi** are 7th semester students in the English Education Department, Faculty of Teacher Training and Education, Universitas Nahdlatul Ulama Surabaya, Indonesia.

REFERENCES

Anggrisia, N. F., & Billah, A. Z. M. (2024). Effective morphological awareness instruction for enhancing vocabulary acquisition in EFL students: A study at AKM Islamic Senior High School. *EXPOSURE: Jurnal Pendidikan Bahasa Inggris, 13*(2), 468–482. https://doi.org/10.26618/exposure.v13i2.15346

Asaad, H. Q. M. (2024). The role of morphological awareness in L2 postgraduates' academic writing: Is vocabulary knowledge a mediating variable? *Cogent Education, 11*(1), Article 2327787. https://doi.org/10.1080/2331186X.2024.2327787

Bauer, L., & Nation, P. (1993). Word families. *International Journal of Lexicography, 6*(4), 253–279. https://doi.org/10.1093/ijl/6.4.253

Bowers, P. N., Kirby, J. R., & Deacon, S. H. (2010). The effects of morphological instruction on literacy skills: A systematic review of the literature. *Review of Educational Research, 80*(2), 144–179. https://doi.org/10.3102/0034654309359353

Carlisle, J. F. (1995). Morphological awareness and early reading achievement. In L. B. Feldman (Ed.), *Morphological aspects of language processing* (pp. 189–209). Lawrence Erlbaum.

Carlisle, J. F. (2010). Effects of instruction in morphological awareness on literacy achievement: An integrative review. *Reading Research Quarterly, 45*(4), 464–487. https://doi.org/10.1598/RRQ.45.4.5

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

Colenbrander, D., von Hagen, A., Kohnen, S., Wegener, S., Ko, K., Beyersmann, E., Behzadnia, A., Parrila, R., & Castles, A. (2024). The effects of morphological instruction on literacy outcomes for children in English-speaking countries: A systematic review and meta-analysis. *Educational Psychology Review, 36*, Article 119. https://doi.org/10.1007/s10648-024-09953-3

Creswell, J. W. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (4th ed.). Pearson.

Ellis, R. (2008). *The study of second language acquisition* (2nd ed.). Oxford University Press.

Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. *College Composition and Communication, 32*(4), 365–387. https://doi.org/10.2307/356600

Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2012). *How to design and evaluate research in education* (8th ed.). McGraw-Hill.

Fritz, C. O., Morris, P. E., & Richler, J. J. (2012). Effect size estimates: Current use, calculations, and interpretation. *Journal of Experimental Psychology: General, 141*(1), 2–18. https://doi.org/10.1037/a0024338

Goodwin, A. P., & Ahn, S. (2013). A meta-analysis of morphological interventions in English: Effects on literacy outcomes for school-age children. *Scientific Studies of Reading, 17*(4), 257–285. https://doi.org/10.1080/10888438.2012.689791

Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics, 66*(1), 64–74. https://doi.org/10.1119/1.18809

Hyland, K. (2019). *Second language writing* (2nd ed.). Cambridge University Press.

Jacobs, H. L., Zinkgraf, S. A., Wormuth, D. R., Hartfiel, V. F., & Hughey, J. B. (1981). *Testing ESL composition: A practical approach*. Newbury House.

Khamraeva Shavkatovna, D., Hashim, H., & Jamaludin, K. A. (2024). Mobile learning for transforming English writing skills: A systematic review (2013–2023). *International Journal of Interactive Mobile Technologies, 18*(19), 4–17. https://doi.org/10.3991/ijim.v18i19.47729

Khan, A. (2024). Transformative impacts of explicit morphological instruction on irregular inflectional acquisition in Saudi EFL learners. *Humanities and Social Sciences Communications, 11*(1), 1–12. https://doi.org/10.1057/s41599-024-03610-4

Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Pergamon.

Kukulska-Hulme, A., & Viberg, O. (2018). Mobile collaborative language learning: State of the art. *British Journal of Educational Technology, 49*(2), 207–218. https://doi.org/10.1111/bjet.12580

Kuo, L.-J., & Anderson, R. C. (2006). Morphological awareness and learning to read: A cross-language perspective. *Educational Psychologist, 41*(3), 161–180. https://doi.org/10.1207/s15326985ep4103_3

Lightbown, P. M., & Spada, N. (2021). *How languages are learned* (5th ed.). Oxford University Press.

Nagy, W., & Townsend, D. (2012). Words as tools: Learning academic vocabulary as language acquisition. *Reading Research Quarterly, 47*(1), 91–108. https://doi.org/10.1002/RRQ.011

Nation, I. S. P. (2022). *Learning vocabulary in another language* (3rd ed.). Cambridge University Press.

Plag, I. (2018). *Word-formation in English* (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. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. M. Gass & C. G. Madden (Eds.), *Input in second language acquisition* (pp. 235–253). Newbury House.

Wahid, R., & Farooq, O. (2019). The influence of derivational and inflectional morphological awareness on the writing of undergraduate EFL students: An empirical study. *Arab World English Journal, 10*(1), 242–258. https://doi.org/10.24093/awej/vol10no1.21

Wang, T., & Zhang, H. (2023). Examining the dimensionality of morphological knowledge and morphological awareness and their effects on second language vocabulary knowledge. *Frontiers in Psychology, 14*, Article 1207854. https://doi.org/10.3389/fpsyg.2023.1207854

Zhou, Y., & Zhou, M. (2026). A meta-analysis on mobile-assisted vocabulary learning: Do mobile applications help? *ReCALL, 38*(1), 75–93. https://doi.org/10.1017/S0958344025100335