Korean Journal of English Language and Linguistics, **Vol 25, September 2025, pp. 1335-1352**
DOI: 10.15738/kjell.25..202509.1335

KOREAN JOURNAL OF ENGLISH LANGUAGE AND LINGUISTICS

ISSN: 1598-1398 / e-ISSN 2586-7474

http://journal.kasell.or.kr

Copying Words: A Double-Edged Sword for Second Language Vocabulary Learning?*

Yechan Lee · Donghyun Kim (Kyungpook National University)

This is an open-access article distributed under the terms of the Creative Commons License, which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Received: July 18, 2025
Revised: September 11, 2025
Accepted: September 16, 2025

Lee, Yechan (First author)
PhD student, Department of English Language and Literature
Kyungpook National University
Email: dldpcks3082@gmail.com

Kim, Donghyun (Corresponding author)
Assistant professor, Department of English Language and Literature
Kyungpook National University
Email: donghyun@knu.ac.kr

*This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2023S1A5A8078806).

ABSTRACT

Lee, Yechan and Donghyun Kim. 2025. Copying words: A double-edged sword for second language vocabulary learning? *Korean Journal of English Language and Linguistics* 25, 1335-1352.

The purpose of this study is to examine whether repetitive writing, often favored by learners as a method of vocabulary memorization, facilitates second language vocabulary learning under supportive conditions. However, it is still unclear whether copying helps vocabulary learning when combined with helpful learning conditions, such as learning words in context and practicing them several times. This study compared copying and observation under controlled conditions that included clear example sentences, repeated exposure, and a follow-up meaning test. Forty participants were divided evenly into two groups and learned target words through either copying or observation. Both groups were exposed to the same example sentences, the same number of repetitions, and the same follow-up meaning test. Mixed-effects regression models showed that the copying group had a slight advantage in the delayed meaning recall test conducted one week after learning, while no significant differences were found between groups in form recognition or form recall. This suggests that copying may help support long-term meaning retention, but overall, observation was just as effective across most tasks. Receptive vocabulary size and working memory did not play a major role in predicting performance. Copying in context may provide temporary benefits for the initial encoding of word form and meaning, but its long-term learning effects may be task-dependent and limited.

KEYWORDS

vocabulary learning, copying strategy, context-based learning, second language acquisition, form recognition, meaning recall, form recall

© 2025 KASELL All rights reserved

1. Introduction

In Korean secondary schools and private institutes, students are often asked to handwrite vocabulary items many times with minimal guidance (Lee 2019). A classroom study with middle-school learners found that using new words in sentences benefited higher-proficiency students, whereas repetitive writing yielded larger gains for lower-proficiency students; despite high satisfaction with the latter, over 80% of students reported a need for explicit strategy instruction (Lee 2009). These findings suggest that handwriting practice is not uniformly ineffective, but that its value may depend on learner profile and on how the activity is implemented. This handwriting practice resembles what Elgort et al. (2018) refer to as "COPYING" (hereafter in small caps), where learners reproduce a given word by visually referring to its form. However, in the Korean educational context, this practice has often been adapted into an extreme form in which a single word is written as many as ten or even twenty times. One concern is the tendency to have students engage in such rote COPYING without sufficient attention to the word's meaning or contextual usage, leading them to focus solely on memorizing the spelling rather than understanding the word. Some learners may continue to use this approach into adulthood. The effectiveness of this COPYING method can vary depending on whether context is provided.

Adult learners, however, typically encounter unfamiliar words in meaningful contexts—university readings, films, or everyday input—where repetition can be embedded in context and linked to comprehension goals. We propose a more structured vocabulary learning framework by emphasizing repeated practice embedded within meaningful contexts, rather than relying on decontextualized repetition. Accordingly, the present study examines how effective COPYING is as a vocabulary learning strategy for adult learners of English. Participants in the COPYING group (experimental group) repeatedly wrote unfamiliar English words, while those in the "OBSERVATION" group (control group; hereafter in small caps) viewed the same words without writing. The relative effectiveness of the two methods was compared. Beyond analyzing differences in learning conditions between the two groups, the study also considered individual differences, focusing on vocabulary size and working memory capacity. All participants took receptive vocabulary size and working memory tests, and their potential roles in vocabulary learning were examined. The learning outcomes were assessed using a range of vocabulary tasks designed to capture different dimensions of lexical knowledge.

2. Background Literature

Vocabulary learning can be approached in two ways, depending on whether the learner is consciously aware of the learning process: intentional learning and incidental learning. Intentional learning means that learners actively try to remember new words, while incidental learning happens naturally, without conscious effort. Inferring the meaning of unfamiliar words from context may be considered incidental learning; however, when exposure is limited or the clues in the sentence are not connected to the word, minimal or no learning occurs (Webb 2008). Webb (2020) also found that when learners are unaware they are learning, it becomes much harder to retain new words. Meaningful and engaging contexts help second language (hereafter "L2") learners focus their attention on new vocabulary (Schmitt 2008). When the writing system of the L2 is very different from that of the learner's L1, word recognition can become more challenging (Hamada and Koda 2008). Therefore, context-based learning is more effective when it incorporates multiple beneficial features, such as repeated exposure and clear example sentences. Building on this background, the present study explores both the strengths and limitations of previous research on COPYING and evaluates copying within a context-based framework to clarify when it supports learning.

© 2025 KASELL All rights reserved

2.1 Copying with Context-Based Word Learning

Several studies have reported positive findings regarding COPYING (e.g. Candry et al. 2018, Elgort et al. 2018, Yoo and Yoon 2019). All of these studies applied context-based approaches to vocabulary learning. In a previous study, Elgort et al. (2018) used the term COPYING to compare a context-based COPYING condition with a meaning-generation task. Their results showed that the COPYING group performed better than the meaning-generation group in both form recall and meaning recall. The study employed immediate post-tests involving recall tasks on spelling and meaning, followed by a delayed test one day later using a lexical decision task to examine the relationship between COPYING and individual differences in terms of accuracy and reaction times. Learners in the COPYING group showed higher accuracy, supporting the idea that COPYING can be an effective learning strategy.

Similarly, Yoo and Yoon (2019) compared a COPYING condition and a meaning inferencing condition within a context-based learning setting. Both conditions were controlled for exposure time by allocating 30 seconds per word. In both immediate and delayed meaning recall tasks, the COPYING group consistently outperformed the inferencing group in accuracy. These findings suggest that COPYING may support vocabulary learning by reinforcing the connection between word form and meaning. Researchers also observed that some participants in the meaning inferencing group misunderstood word meanings, which led to errors. In the form recall task, the COPYING condition led to higher accuracy on the immediate test; however, this positive effect did not persist in the delayed post-test. These results suggest that COPYING may better support long-term retention of word meaning than of word form.

Candry et al. (2018) compared three learning conditions—repeated writing, repeated speaking, and simple viewing (control)—among Dutch-speaking learners of German as an L2. Participants completed form recall, meaning recall, and form recognition tasks immediately after learning, and the same tasks were repeated in a delayed post-test one week later. In the immediate test, participants in the repeated writing group had the highest form recall scores. However, in the delayed test, differences between groups were minimal. Furthermore, unlike previous studies, the writing condition did not show a clear advantage in meaning recall. These findings led the authors to conclude that the benefits of writing are primarily limited to short-term learning effects.

2.2 Copying with Word-Meaning Pairs

Some studies take a critical view of COPYING (Barcroft 2006, Candry et al. 2020, Kakihana 2024, Webb and Piasecki 2018). A common feature of these studies is that they used learning tasks without sentence-level contexts, instead relying on simpler formats such as word pairs or word–picture associations. For instance, Barcroft (2006) compared a COPYING condition and a visual presentation condition in native English speakers learning Spanish. Participants were presented with Spanish words paired with corresponding images, and those in the writing condition wrote each word by hand once, while those in the visual condition simply viewed the words. The 24 target words were divided into two lists (items 1–12 and 13–24), and each pair was presented twice, with a 6-second presentation time for each. In both the immediate post-test and the delayed post-test conducted two days later, participants in the visual condition recalled Spanish spellings more accurately than those in the writing condition. One possible explanation is that, in the COPYING condition, learners had only a single opportunity to write each word under time pressure immediately after viewing it, which may have split their attentional resources. Moreover, the limited number of repetitions may not have been sufficient to establish motor memory. Therefore, the lower performance in the writing group may not reflect the act of writing itself, but rather suboptimal learning conditions—specifically, insufficient focus and repetition.

© 2025 KASELL All rights reserved

Similarly, Webb and Piasecki (2018) conducted an experiment involving intermediate to advanced EFL learners. Participants were presented with English word–picture pairs and were assigned to one of three conditions: writing the word for six seconds, viewing the word for six seconds without writing, or writing the word without a time limit. In the subsequent tests, including a multiple-choice form recognition task and form production tasks, no significant differences were found between the six-second writing condition and the non-writing condition. The time constraint imposed in the six-second writing condition may have hindered effective encoding by inducing cognitive load. Kakihana (2024) also investigated COPYING using L2–L1 word pairs. Participants were asked to memorize both the spelling and meaning of each L2 target word while viewing its corresponding L1 translation. Participants in the writing condition wrote out the target words by hand within a 15-second learning period. Learning was assessed through a multiple-choice form recognition test and a free-response meaning recall test, both of which were repeated one week later in a delayed post-test. Neither the immediate post-test nor the delayed post-test revealed significant group differences in either form recognition or meaning recall. These findings suggest that the COPYING condition did not yield substantial learning gains in either meaning recall or form recognition.

Candry et al. (2020) compared a word writing condition with a simple visual presentation condition in Dutch-speaking English learners aged 15 to 18. Participants learned 15 English nouns across two phases. In Phase 1, all participants were shown each target word along with its Dutch translation for 10 seconds to memorize. In Phase 2, learners in the writing condition were presented with only the English words (without translations) for 15 seconds and instructed to write them repeatedly. In the control condition, participants reviewed the English–Dutch word pairs for 15 seconds, following the same procedure as in Phase 1. In the immediate form recall test, the writing group scored higher than the control group (Writing: 62.0%, Control: 52.9%). However, one week later, the writing group's accuracy dropped sharply, while the control group's performance remained more stable (Writing: 27.1%, Control: 36.6%). In the meaning recall task, the control condition consistently outperformed the writing condition across both test times (Immediate: Writing 79.1%, Control 83.1%; Delayed: Writing 76.0%, Control 79.0%). The results indicated that the group differences were statistically significant. Taken together, these studies suggest that the COPYING strategy is less effective when based on decontextualized word–meaning pairs and does not consistently support long-term retention.

2.3 Individual Differences Factors

This study extends beyond a simple comparison of learning conditions by investigating how individual differences, specifically receptive vocabulary size and working memory capacity, influence vocabulary learning outcomes. It further aims to identify learner profiles for whom the COPYING strategy is particularly effective. More specifically, the study seeks to assess how vocabulary size and working memory modulate the effectiveness of COPYING. According to Daidone (2020), receptive vocabulary size had a greater impact on L2 vocabulary processing than cognitive or perceptual ability. This supports the lexicon-first model, which suggests that learners reconstruct new words more easily by drawing on similarities with already known words. Elgort (2018) also found a strong positive link between receptive vocabulary size and form recognition, showing that learners with larger vocabularies are better at recognizing unfamiliar word forms. However, Candry et al. (2018) found that receptive vocabulary size predicted immediate test scores but not delayed ones. These findings suggest that while vocabulary size helps in short-term learning, it may not reliably predict long-term retention.

Working memory also plays a key role as an individual difference factor. Previous studies based on Baddeley's model of working memory (Li 2021, Ruiz et al. 2021) have reported that learners with greater working memory capacity demonstrate enhanced outcomes in form-focused learning. Elgort (2018) measured working memory

© 2025 KASELL All rights reserved

using the O-span task and found that learners with higher working memory capacity demonstrated approximately 40% greater accuracy in form production under the COPYING condition compared to the meaning inferencing condition. This suggests that working memory functions beyond short-term storage and significantly influences learning outcomes depending on the instructional condition and task type. It also implies that working memory may be a valid predictor of cognitive differences among learners that cannot be explained by vocabulary size alone.

2.4 Methodological Framework of the Present Study

Building on previous research (Candry 2017, Elgort et al. 2018), this study focuses on context-based vocabulary learning and takes a different approach from earlier studies that used word pairs or picture-based materials without context. This approach reflects the fact that adult learners usually encounter unfamiliar words in meaningful contexts, not in isolated word lists or images. In addition, based on Milton and Hopkins (2006), who found that advanced learners rely more on spelling than on sound, this study compares a COPYING condition with an OBSERVATION condition using only visual input, without audio. Several key factors must be considered to ensure the effectiveness of this strategy. First, learning with context is likely more effective than learning with isolated words. The context should focus on meaning and help learners guess and understand word meanings (Webb 2008). Controlling the number of repetitions, rather than the total duration, is essential to avoid variability in learners' engagement and pacing. Excessive repetition (e.g. 10 or 20 times) may lead to cognitive overload and reduce learning outcomes; hence, determining an appropriate number of repetitions that balances effectiveness and fatigue is crucial. Also, if one exposure to a word list is considered a "block," deciding how many blocks to include is another key design choice. As Hui (2020) pointed out, more blocks do not always lead to better learning, and learner attention and fatigue should also be considered. Finally, to keep vocabulary learning from becoming mechanical, tasks should connect both form and meaning in a clear learning sequence.

Vocabulary assessment should go beyond counting correct answers and instead reflect the multidimensional nature of vocabulary knowledge. In L2 learning, acquiring vocabulary usually involves linking the written form of a word with its meaning, known as the form–meaning connection. Functionally, this can be divided into two categories: vocabulary recognition refers to the activation of the learner's L1 meaning upon seeing the foreign word form, whereas vocabulary recall involves retrieving either the form or the meaning in response to a specific cue or context. When applied to task types, recognition involves identifying the correct answer from multiple choices, while recall requires learners to produce the correct word form or meaning without any options. Laufer and Goldstein (2004) proposed a hierarchy of difficulty: meaning recognition < form recognition < meaning recall < form recall. This implies that learners generally find meaning recognition easier than form recognition and perform better on recognition tasks than recall tasks. Recall, in particular, poses a greater challenge for L2 learners than recognition. Therefore, vocabulary assessment should include these distinctions and apply them consistently across both immediate and delayed post-tests. If the COPYING condition yields greater learning gains than the OBSERVATION condition across various dimensions—such as form recognition, meaning recall, and form recall—and these gains are maintained after one week with minimal decline, this would suggest that COPYING, when supported by contextual input, serves as an effective strategy for vocabulary learning.

Taken together, the present study implemented rigorous controls for variables including sentence-level context, presentation order, repetition count, exposure duration, task type, and test timing. Target words were chosen to be unfamiliar to all participants, regardless of group. Under these conditions, the study examines how COPYING and OBSERVATION affect form recognition, meaning recall, and form recall, both immediately and after a one-week delay. The goal is to compare learning gains and retention between the two learning conditions within the same

© 2025 KASELL All rights reserved

controlled environment. The research questions of this study are as follows.

1) In the form recognition task, do learners who study by COPYING demonstrate higher accuracy than those who only observe the words, and is this difference sustained over time?

2) In the meaning recall task, do learners who study by COPYING demonstrate higher accuracy than those who only observe the words, and is this difference sustained over time?

3) In the form recall task, do learners who study by COPYING demonstrate higher accuracy than those who only observe the words, and is this difference sustained over time?

4) How do working memory and vocabulary size (i.e. prior vocabulary knowledge) influence learning outcomes in the form recognition, meaning recall, and form recall tasks?

3. Methods

3.1 Participants

A total of 40 participants (mean age = 22.98, range = 18–32 years, 24 female) took part in the study. They were undergraduate or graduate students at a university in South Korea, native speakers of Korean, and between 18 and 35 years of age. All participants had a TOEIC score of 800 or higher. Although the mean TOEIC scores of the two groups were not identical, the score was used solely as a recruitment criterion to ensure a minimum proficiency threshold. More relevant to this study, however, is that a pre-test vocabulary assessment confirmed that participants in both groups were generally unfamiliar with the target words prior to learning. They were randomly assigned to one of two groups: the COPYING group or the OBSERVATION group. Table 1 summarizes the demographic characteristics of the overall sample and each group.

Table 1. Summary of Group Size, Mean Age, and Mean TOEIC Score by Experimental Condition

Group	*N*	Mean Age	Mean TOEIC Score
Total	40	22.98	889.95
COPYING	20	23.05	904.65
OBSERVATION	20	22.90	875.25

Note: *N* = number of participants. TOEIC = Test of English for International Communication.

3.2 Stimuli

This study used 16 low-frequency English words as target vocabulary items. Words were selected based on the Corpus of Contemporary American English (COCA), with a cut-off frequency of 2,000 occurrences or fewer. This was intended to ensure that the vocabulary items were relatively unfamiliar to Korean learners of English. For the form recognition and form recall tasks, only words containing between five and ten alphabetic letters were included. For the meaning recall task, only words that could be clearly translated into Korean were used. The selected words were not commonly found on standardized English tests such as the CSAT or TOEIC. For example, “ladle” and “offal” were adopted from Candry (2017) due to their semantic clarity and ease of translation. In contrast, words

© 2025 KASELL All rights reserved

such as “fontina,” “gable,” and “busser” were excluded due to potential cultural barriers in conceptual understanding. Conversely, words like “hubris,” “zealot,” “proclivity,” and “paucity” were included, as their meanings remain inferable despite low familiarity (Chen et al. 2024). Remaining items were finalized based on COCA frequency and verified for clarity using Korean translations in the Naver Dictionary. All target words were presented in parentheses at the top center of each sentence, allowing learners to read the entire sentence containing the target word. The complete list of target words can be found in Appendix A. Sentences were designed to provide enough contextual information for learners to infer word meaning. Reliable sources such as the Naver English Dictionary, Collins Dictionary, and Oxford English Dictionary were consulted to select example sentences with clear meanings. If a sentence was excessively long, it was shortened to reduce cognitive load during learning. Three native English speakers reviewed the final sentence selections to ensure clarity. For example, learners encountered sentences such as “He pursued his goals with great (ardor),” where the target word appears in parentheses, embedded in context to support meaning inference.

3.3 Procedure

The experiment consisted of three phases: participant recruitment, learning, and testing. Participants were recruited via the university platform *Everytime* and through word-of-mouth. They registered by completing a Google Form. Based on their TOEIC scores and overseas residency experience, eligible participants were individually contacted to schedule their participation. Selected participants confirmed their participation in both sessions via email and text message. The experiment was conducted in a quiet laboratory setting. Before the first session, participants received a printed list of the target vocabulary items and completed a pre-test vocabulary assessment. They were instructed to circle each word they recognized, and their responses were immediately verified by the researcher. Responses were categorized as “known,” “partially known,” or “unknown.” The pre-test results indicated that only two participants reported knowing the meanings of one or two target words, confirming that the vocabulary items were generally unfamiliar to the vast majority of participants at the outset of the experiment. On the day of the first experiment, participants first completed the LexTALE vocabulary test (Lemhöfer and Broersma 2012) to assess their general English proficiency, followed by the vocabulary learning task. Participants in the COPYING group read contextualized sentences and repeatedly typed the target word. Blocks 1–4 required three repetitions, while blocks 5–8 required two repetitions. After each block, a four-option meaning recognition task was administered. The OBSERVATION group was exposed to the same sentences and blocks but instructed only to read and click through the sentences without typing. This design aimed to encourage learner autonomy. Immediately after the learning phase, three post-tests were administered. The form recognition task involved distinguishing real words from nonwords. The meaning recall task required participants to describe the meaning of a target word presented in parentheses, whereas the form recall task asked them to type the English word using only the first letter as a cue.

The second session was conducted one week later (within 6 to 8 days, including holidays) at the same location and consisted of a delayed post-test and a working memory task. The delayed post-test was administered in the same format as the immediate post-test to assess retention of the learned material. Afterward, participants completed a working memory task consisting of the digit span tasks, which included a total of 24 items: 14 forward span items and 10 backward span items. In the forward span task, participants recalled numbers in the order presented; in the backward span task, they recalled them in reverse. Each task included both practice and test phases. The backward span task is considered cognitively demanding, as it engages not only the phonological loop but also the central executive function (Juffs and Harrington 2011). To enhance engagement and motivation,

© 2025 KASELL All rights reserved

participants received compensation of 25,000 KRW upon completing the experiment. The overall experimental procedure is illustrated below.

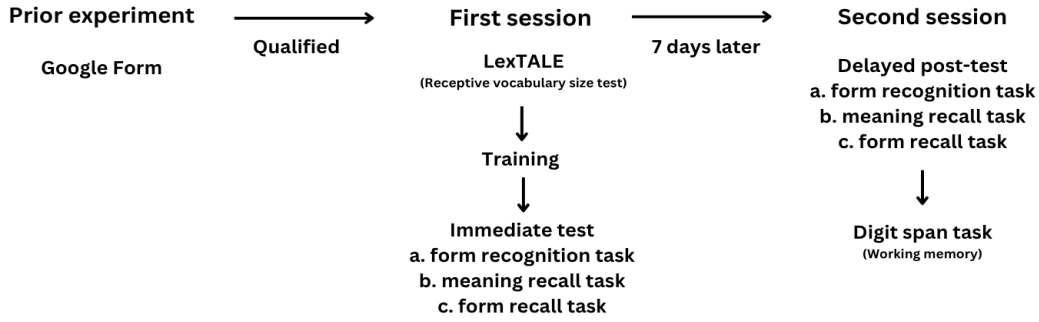

Figure 1. Experiment Procedure Overview

3.4 Scoring

To quantitatively assess participants' vocabulary learning outcomes and individual difference variables, scores were calculated from three vocabulary tasks and two individual difference measures. The vocabulary tasks consisted of form recognition, meaning recall, and form recall. The form recognition task involved multiple-choice questions and was scored using a binary method: correct answers received a score of 1, and incorrect answers received a score of 0. Accuracy was calculated as the proportion of correct responses. In the meaning recall task, participants' responses were evaluated based on whether they conveyed the correct meaning, a partially related meaning, or an unrelated meaning. Three independent evaluators checked each response to determine whether it accurately conveyed the target meaning. A score of 1 was assigned if the response clearly conveyed the meaning and closely matched the Korean definition provided during the learning process. Semantically related but imprecise responses were scored as 0.5, and unrelated or incorrect responses were scored as 0. Only the meaning recall task was evaluated by all three evaluators independently, and final scores were determined through consensus. The total raw score was calculated based on 16 items and converted into a proportion. In the form recall task, responses were scored as follows: correctly spelled words received 1 point, whereas blank or misspelled responses received 0. The total score for the 16 items was also converted into a proportion and used as the final score.

In addition, participants' L2 proficiency was assessed using the LexTALE test, and their cognitive ability was evaluated through a working memory task. LexTALE scores were calculated according to the formula proposed by its original developers: ((number of correct words / 40 × 100) + (number of correct nonwords / 20 × 100)) / 2. Working memory scores were evaluated following the criteria outlined by Conway et al. (2005). Although their framework informed the overall scoring approach, a graded scoring system was adopted in place of binary scoring to more accurately reflect varying task difficulty. For example, participants who correctly recalled digit strings of seven or more digits in the forward task or five or more digits in the backward task were awarded partial credit (0.5 points). Final working memory scores were calculated by summing the number of correct responses and the additional points awarded for difficulty level.

© 2025 KASELL All rights reserved

3.5 Data Analysis

During the learning phase, the average response time for the four-option multiple-choice meaning recognition task, presented after each target word, gradually decreased across blocks, suggesting that participants were actively engaged in the task. Appendix B provides a summary of participants' accuracy and response times across blocks during the learning phase. In the form recognition task administered immediately after learning, no participants showed signs of unnatural behavior, such as repeatedly pressing the same key or producing overly consistent answer patterns. Accordingly, participants with lower scores on the delayed post-test were retained in the analysis, as no evidence of artificial responding was observed. All statistical analyses were performed using R version 4.4.1 (R Core Team 2024). The primary analyses employed mixed-effects models using the *lme4* package (Bates et al. 2015). Participants' vocabulary learning performance was evaluated using trial-level mixed-effects models constructed separately for the three tasks: form recognition, meaning recall, and form recall. For the form recognition and form recall tasks, which involved binary responses (correct/incorrect), generalized linear mixed-effects models (GLMMs) were used. In contrast, the meaning recall task, which involved continuous scores (0, 0.5, or 1), was analyzed using a linear mixed-effects model (LMM). All models included random intercepts for both participants and words to account for participant and word specific variability of the data. Contrast coding was applied to categorical variables. For the test variable, which had three time points (pre, immediate, delayed), backward difference coding (*contr.sdif*) was used. For the form recognition and form recall tasks, both group and test variables were coded using sum coding (−0.5, 0.5). All coding variables were centered using the *rescale()* function from the *arm* package (Gelman and Su 2024). Continuous individual difference variables, including working memory scores and LexTALE scores, were also standardized using the same function to facilitate interpretability. Model results were summarized using the *model_parameters()* function from the parameters package. Accuracy results were visualized using *ggplot2* (Wickham 2016), with jittered points for individual participants and group-level trend lines presented to visually compare learning effects across conditions.

4. Results

Before examining each task, it is important to note that participants initially did not know the target words, yet achieved proportion correct scores above 0.6 across all tasks and test sessions. This suggests that vocabulary learning occurred successfully under both the COPYING and OBSERVATION conditions.

4.1 Form Recognition

Table 2. Summary of Logistic Mixed-Effects Model for Form Recognition

Parameter	*β*	*SE*	*CI* Low	*CI* High	*p*
Intercept	2.327	0.222	1.893	2.761	< 0.001
Group (COPYING vs. OBSERVATION)	0.428	0.256	−0.074	0.930	0.094
Test (Delayed vs. Immediate)	−0.485	0.125	−0.792	−0.240	< 0.001
Working memory (std)	0.276	0.249	−0.212	0.764	0.268
Vocabulary size (std)	−0.051	0.258	−0.558	0.455	0.842
Group × Test	0.129	0.254	−0.368	0.626	0.611
Test × Working memory	−0.025	0.249	−0.513	0.462	0.919
Test × Vocabulary size	0.080	0.252	−0.415	0.574	0.752

Note: Model coefficient estimates (*β*), standard errors (*SE*), confidence intervals (*CI*), and *p*-values.

© 2025 KASELL All rights reserved

As shown in Table 2, no statistically significant difference in form recognition accuracy was found between the COPYING group and the OBSERVATION group. However, there was a general trend of decreased accuracy in the delayed post-test conducted one week after learning. In this task, neither working memory nor receptive vocabulary size significantly predicted accuracy. Moreover, no interaction effect was observed between test session (immediate vs. delayed) and working memory.

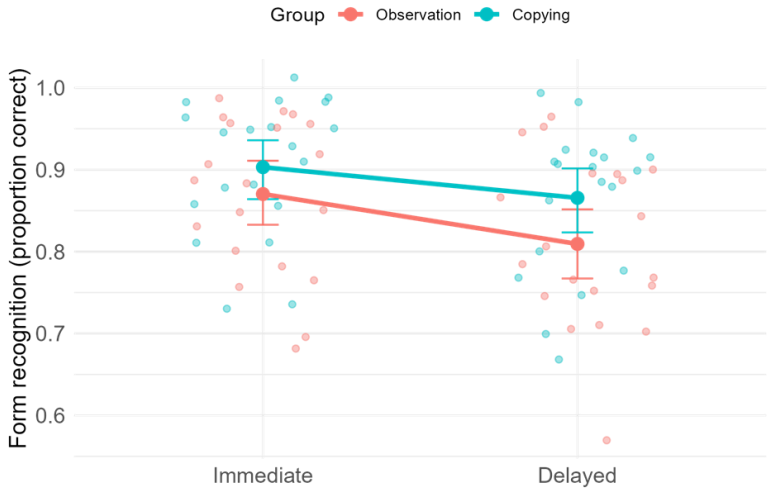

Figure 2. Proportion Correct in Form Recognition Across Test Sessions

Figure 2 illustrates individual accuracy scores, group means, and 95% confidence intervals for the form recognition task administered both immediately after learning and one week later. Both the COPYING group (M = 0.91, SD = 0.30) and the OBSERVATION group (M = 0.87, SD = 0.34) showed similarly high accuracy levels immediately after learning, with no statistically significant difference between the groups. In the delayed post-test, the COPYING group (M = 0.87, SD = 0.34) and the OBSERVATION group (M = 0.81, SD = 0.39) also did not differ significantly in accuracy, and both groups maintained a high level of form recognition overall.

4.2 Meaning Recall

Table 3. Summary of Linear Mixed-Effects Model for Meaning Recall

Parameter	β	SE	CI Low	CI High	p
(Intercept)	0.564	0.022	0.521	0.606	< 0.001
Group (COPYING vs. OBSERVATION)	0.015	0.028	−0.040	0.071	0.592
Test 1 (Immediate − Pre)	0.914	0.015	0.885	0.944	< 0.001
Test 2 (Delayed − Immediate)	−0.152	0.015	−0.181	−0.123	< 0.001
Working memory (std)	−0.028	0.028	−0.082	0.026	0.309
Vocabulary size (std)	−0.010	0.029	−0.066	0.046	0.722
Group × Test 1	0.003	0.031	−0.058	0.063	0.934
Group × Test 2	0.066	0.031	0.005	0.126	0.035
Test 1 × Working memory	−0.047	0.030	−0.106	0.012	0.118
Test 2 × Working memory	−0.003	0.030	−0.062	0.057	0.929
Test 1 × Vocabulary size	−0.031	0.031	−0.093	0.030	0.315
Test 2 × Vocabulary size	0.019	0.031	−0.043	0.080	0.550

Note: Model coefficient estimates (β), standard errors (SE), confidence intervals (CI), and p-values.

© 2025 KASELL All rights reserved

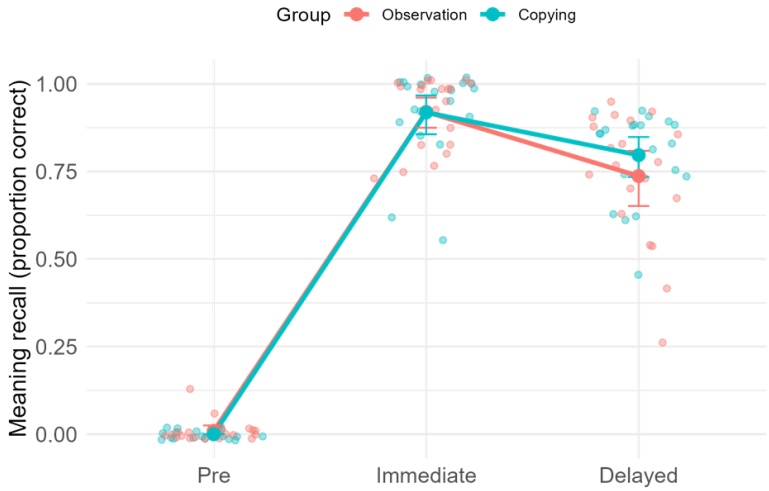

Figure 3. Proportion Correct in Meaning Recall Across Test Sessions

Figure 3 illustrates that participants who initially did not know the meanings of the target words achieved high levels of meaning recall accuracy immediately after learning (M = 0.92, SD = 0.27). Both the COPYING and OBSERVATION groups showed identical average accuracy at this time point. However, in the delayed post-test administered one week later, the COPYING group (M = 0.80, SD = 0.40) retained significantly higher accuracy than the OBSERVATION group (M = 0.74, SD = 0.43). These findings suggest that copying-based learning may be more effective for supporting the retention of semantic knowledge.

4.3 Form Recall

Table 4. Summary of Logistic Mixed-Effects Model for Form Recall

Parameter	β	SE	CI Low	CI High	p
(Intercept)	1.015	0.293	0.440	1.589	< 0.001
Group (COPYING vs. OBSERVATION)	0.605	0.387	−0.155	1.364	0.119
Test (Delayed vs. Immediate)	0.131	0.138	−0.141	0.402	0.345
Working memory (std)	0.055	0.378	−0.686	0.797	0.884
Vocabulary size (std)	0.144	0.389	−0.618	0.906	0.711
Group × Test	−0.488	0.289	−1.055	0.079	0.092
Group × Working memory	0.120	0.281	−0.431	0.672	0.669
Test × Vocabulary size	0.112	0.288	−0.453	0.677	0.698

Note: Model coefficient estimates (β), standard errors (SE), confidence intervals (CI), and p-values.

According to the results presented in Table 4, no statistically significant difference was found between the COPYING and OBSERVATION groups in the immediate post-test for the form recall task. Similarly, no significant group difference in accuracy was observed in the delayed post-test administered one week later. Additionally, individual difference variables such as working memory and vocabulary size did not significantly predict form recall accuracy.

© 2025 KASELL All rights reserved

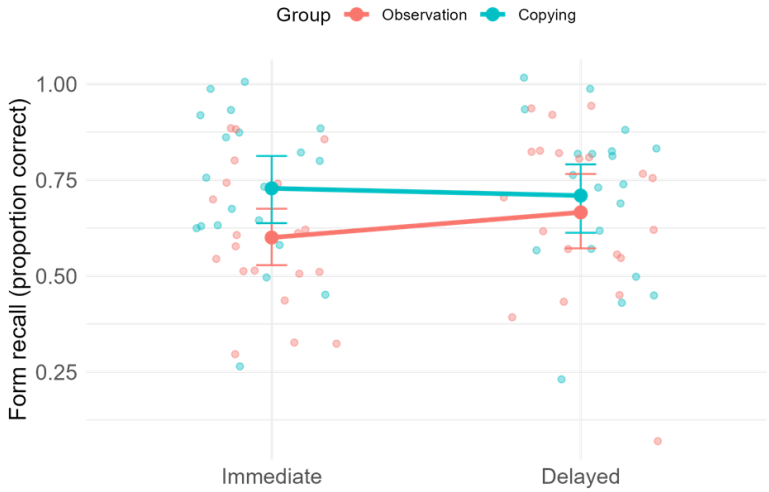

Figure 4. Proportion Correct in Form Recall Across Test Sessions

Figure 4 shows that in the immediate post-test, the COPYING group (M = 0.73, SD = 0.45) achieved higher accuracy in the form recall task compared to the OBSERVATION group (M = 0.60, SD = 0.49). However, in the delayed post-test, the COPYING group's accuracy slightly declined (M = 0.71, SD = 0.46), whereas the OBSERVATION group showed a slight increase in accuracy (M = 0.67, SD = 0.48).

5. Discussion

In South Korea, rote memorization strategies that involve COPYING—often referred to as *ggamji*, *bbakbbaki*, or *ppaekppaegi*—have long been used in educational settings. This practice typically requires students to write unfamiliar words 10 to 20 times to aid memorization but may result in students writing words incorrectly. A similar approach has been documented in Japanese classrooms (Kakihana 2024), where it has been criticized for creating an illusion of diligence while requiring substantial time and effort with little learning efficiency. Many studies have noted the limitations of such mechanical repetition for retention of L2 word forms and meanings. In real learning contexts, learners encountering unfamiliar words in context often use dictionaries or AI-based tools to access meanings. The present study compared the effects of COPYING and OBSERVATION on three aspects of vocabulary learning: form recognition, meaning recall, and form recall. Overall, there were no statistically significant differences between the two groups in form recognition and form recall accuracy across both immediate and delayed post-tests, although both groups maintained relatively high levels of performance. In contrast, a significant difference emerged in the delayed meaning recall task, where the COPYING group showed significantly higher accuracy than the OBSERVATION group one week after learning, suggesting that COPYING may support the retention of semantic knowledge. Neither receptive vocabulary size (LexTALE) nor working memory significantly predicted performance across any of the tasks. The null effects can be explained by several design constraints. First, the participant sample was deliberately homogeneous, with all learners scoring 800 or higher on the TOEIC, which restricted variance in proficiency. Second, the recognition tasks showed high mean accuracy, creating a possible ceiling effect that could obscure individual differences. Third, the digit-span task included a limited number of items, resulting in a narrow score range. Moreover, digit-span tasks primarily assess short-term storage

© 2025 KASELL All rights reserved

and manipulation rather than the context-integration processes emphasized in this study. Taken together, these factors may have reduced the likelihood of detecting genuine contributions of vocabulary size or working memory. Future research should therefore recruit learners across a wider proficiency range, employ more extensive working-memory measures, and adjust task difficulty to better capture individual-difference effects.

Yoo and Yoon (2019) compared a COPYING condition with a meaning inference condition among Korean high school students. They found that although both groups showed immediate learning gains in meaning recall and form recall, the COPYING condition exhibited a decline in performance in the delayed post-test. In that study, COPYING was done without L1 annotations, and the effect may have stemmed not from COPYING itself but from the cognitive demands of inference and the lack of feedback in the comparison group. For first-year high school students, contextual inference may impose excessive cognitive load, and without proper feedback, incorrect guesses can lead to entrenched misunderstandings. Thus, the observed decline in the COPYING group may reflect a design artifact rather than a fundamental flaw in COPYING as a strategy. In fact, in context-based learning environments, COPYING may facilitate form recognition to some extent. In contrast to high school learners in Yoo and Yoon (2019), who showed greater immediate benefits from word writing, the present study found that among advanced adult learners, repeated contextual exposure alone was largely sufficient to support form learning. Although the COPYING group demonstrated slightly higher accuracy, the difference was not decisive. The finding that all groups retained over 80% accuracy in form recognition one week later underscores that, for learners who already possess a substantial proficiency level, the presence of meaningful context is more critical than the act of COPYING itself.

Previous studies have reported that COPYING words in isolation, without contextual support, does not significantly enhance semantic learning (Candry et al. 2020, Kakihana 2024). However, in the present study, when advanced L2 learners engaged in COPYING target words that were embedded in meaningful sentence contexts, they tended to retain the meanings of those words more effectively over time. This finding aligns with Candry et al. (2018), suggesting that the act of writing may involve more than rote orthographic repetition. Rather, it may involve deeper cognitive processing, prompting learners to attend to the surrounding context and internally activate word meanings during the COPYING task. This is also consistent with Yoo and Yoon (2019), who found that contextual word learning can benefit from more active processing strategies. These results suggest that, as a learning strategy, COPYING may be more effective than passive OBSERVATION for long-term semantic retention. Specifically, learners may have engaged in subvocalization—silent verbal rehearsal during reading—which could have internally activated L1 meanings. According to Perrone-Bertolotti et al. (2012), subvocalization supports semantic access, comprehension, and memory retention while also reducing cognitive load. Similarly, Vilhauer (2017) demonstrated that many individuals process written language via their inner voice, even when reading silently. Although this study did not measure the depth of contextual reading qualitatively, it is plausible that highly proficient EFL learners spontaneously used inner speech while reading and COPYING. Taken together, these observations suggest that the COPYING task may have involved not only orthographic memorization but also a process of internalizing meaning through context-driven semantic reactivation.

Given that COPYING may engage motor memory, it was initially expected that COPYING target words would yield greater benefits for form recall compared to simple OBSERVATION. Although form recall accuracy was relatively high, consistent with Candry et al. (2018), the findings contrast with those of Yoo and Yoon (2019). One possible explanation is that COPYING was systematically limited in the present study, whereas Yoo and Yoon allowed learners to copy as many times as they wished. This suggests that externally constrained repetition may weaken the immediate benefits of COPYING on form recall. Interestingly, the practice effect—improved performance across repeated testing—was more pronounced in the OBSERVATION group than in the COPYING group during the delayed

© 2025 KASELL All rights reserved

post-test. This pattern partially mirrors Candry et al. (2020), where the COPYING group outperformed the OBSERVATION group in the immediate test, but the reverse was true in the delayed test. Although the reversal in accuracy was not statistically significant in the present study, the improved performance in the OBSERVATION group after a one-week interval deserves attention. This pattern is consistent with a cognitive-load account. Contrary to expectations that COPYING would promote dual encoding of form and meaning, retention declined, likely because attention was divided between motor execution and contextual processing. The repeated, spaced exposures across blocks may have supported consolidation without the interference caused by immediate production. Under these conditions, observation approximates a low-load, intentional encoding routine that favors later access to form. By contrast, participants in the OBSERVATION condition—who repeatedly viewed words in context without producing them—may have encoded word forms more effectively without cognitive overload. This finding suggests that learning through OBSERVATION within meaningful contexts may facilitate the integration of both form and meaning with less cognitive strain. In other words, observation-based learning, despite the lack of overt production, may have been more advantageous for form recognition. A similar finding was reported by Kapnoula and Samuel (2023), who examined the encoding of novel words using auditory stimuli. They compared conditions in which participants immediately repeated the words, delayed their repetition by 2 or 4 seconds, or merely listened without speaking. Although their study involved auditory rather than visual input, a critical distinction emerged between active production and passive encoding. They found that immediate repetition led to the lowest learning outcomes, while performance improved with longer delays, and was highest when no production occurred. These findings suggest that immediate production can interfere with encoding, whereas passive exposure may help integrate new words into the mental lexicon. As Barcroft (2006) noted, writing new words can impose additional cognitive demands, reducing the resources available for effective meaning encoding.

These findings align with the pattern observed in the current study's form recall task. One week after learning, overall recall performance declined, and notably, learners in the COPYING condition showed no clear improvement. In contrast, learners in the OBSERVATION condition seemed to develop more accurate orthographic representations by using cue-recall strategies from contextual information. This suggests that integrating form and meaning in sentence context may support stable lexical consolidation, while COPYING may have disrupted this process. While some participants may engage in inner reading—internally rehearsing the word during COPYING —this cannot be entirely ruled out. However, individual differences in working memory capacity were minimal in this study, and working memory was not a significant predictor of performance in the form recall task. Therefore, it is more reasonable to interpret the observed differences as stemming from the nature of the learning activity rather than from learners' cognitive capacities. These results suggest that learning methods involving multiple sensory demands, such as writing, may induce excessive cognitive load for learners. Even though participants in the present study were advanced L2 learners with relatively high proficiency, there were no significant differences between the COPYING and OBSERVATION groups in most tasks, except for meaning recall, either immediately after learning or after a one-week delay. This implies that in context-based vocabulary learning, the cognitive process of understanding the sentence and integrating form and meaning may be more crucial than the act of writing itself. This interpretation resonates with the findings of Kapnoula and Samuel (2023), who, despite modality differences, also demonstrated that overt production can sometimes hinder learning outcomes.

While this study offers important insights by comparing COPYING and OBSERVATION under context-based vocabulary learning, several limitations remain, along with directions for future research. First, the working memory measure employed in this study used a limited number of digit span items, resulting in a restricted score range. This may have constrained the precision with which individual differences in working memory capacity could be captured. Future studies should consider expanding the number of items and diversifying difficulty levels

© 2025 KASELL All rights reserved

to allow for more refined analyses of individual difference variables. Second, the order of the tests—form recognition, meaning recall, and form recall—was fixed across all participants. This fixed order may have led to practice effects in the form recall task. To mitigate such order effects, future studies should consider counterbalancing or randomizing the sequence of tasks. Finally, incorporating qualitative approaches could enhance understanding of participants' contextual comprehension and vocabulary learning strategies. For example, follow-up interviews or questionnaires could explore whether participants engaged in inner reading, how well they understood each sentence, how they perceived the target word's meaning, and what strategies they used for vocabulary learning. Such data could offer a more nuanced view of how learning conditions relate to learners' strategies during the task.

6. Conclusion

This study compared the effects of two learning conditions—COPYING and OBSERVATION—on form recognition, meaning recall, and form recall in a context-based vocabulary learning environment. It also examined how individual differences in vocabulary size and working memory influenced learning outcomes. The analysis revealed that the COPYING strategy led to relatively better performance in the delayed test of the meaning recall task. However, overall, the OBSERVATION condition maintained a comparable level of learning effectiveness in both form recognition and meaning recall. Neither working memory nor receptive vocabulary size emerged as a significant predictor. These findings suggest that, in foreign language education, instructional designs emphasizing contextualized and cognitively efficient input may be more effective than rote COPYING strategies. They also highlight the need for a tailored approach to vocabulary instruction that considers cognitive load and the use of contextual cues.

References

이아진(Lee, A. J.). 2019. 중학생의 어휘학습을 위해 사용하는 전략에 대한 조사(Investigating Strategies for Middle School Students to Learn Vocabularies). 석사학위논문(Master's thesis). 조선대학교(Chosun University).

Barcroft, J. 2006. Can writing a new word detract from learning it? More negative effects of forced output during vocabulary learning. *Second Language Research* 22(4), 487-497.

Bates, D., M. Mächler, B. M. Bolker, M. and S. Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1), 1-48.

Candry, S., J. Deconinck and J. Eyckmans 2018. Written repetition vs. oral repetition: Which is more conducive to L2 vocabulary learning? *Journal of the European Second Language Association* 2(1).

Candry, S., J. Decloedt and J. Eyckmans. 2020. Comparing the merits of word writing and retrieval practice for L2 vocabulary learning. *System* 89, 102206.

Candry, S., I. Elgort, J. Deconinck and J. Eyckmans. 2017. Word writing vs. meaning inferencing in contextualized L2 vocabulary learning: Assessing the effect of different vocabulary learning strategies. *Canadian Modern Language Review* 73(3), 293-318.

Chen, P. H., R. C. Hulme, L. M. Blott. and J. M. Rodd 2024. The web-based open-access reliable decision on

© 2025 KASELL All rights reserved

synonyms (WORDS) English vocabulary test. *Journal of Cognition* 7(1), 64.

Conway, A. R., M. J. Kane, M. F. Bunting, D. Z. Hambrick, Wilhelm, O. and Engle, R. W. 2005. Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review* 12, 769-786.

Daidone, D. 2020. *How Learners Remember Words in Their Second Language: The Impact of Individual Differences in Perception, Cognitive Abilities, and Vocabulary Size.* Doctoral dissertation, Indiana University.

Daidone, D. and I. Darcy. 2021. Vocabulary size is a key factor in predicting second language lexical encoding accuracy. *Frontiers in Psychology* 12, 688356.

Elgort, I., S. Candry, T. J. Boutorwick, J. Eyckmans. and M. Brysbaert. 2018. Contextual word learning with form-focused and meaning-focused elaboration. *Applied Linguistics* 39(5), 646-667.

Gelman, A. and Y.-S. Su. 2024. arm: Data analysis using regression and multilevel/hierarchical models (R package version 1.14-4) Comprehensive R Archive Network.

Hamada, M. and K. Koda. 2008. Influence of first language orthographic experience on second language decoding and word learning. *Language Learning* 58(1), 1-31.

Hui, B. 2020. Processing variability in intentional and incidental word learning: An extension of Solovyeva and DeKeyser (2018). *Studies in Second Language Acquisition* 42(2), 327-357.

Juffs, A. and M. Harrington. 2011. Aspects of working memory in L2 learning. *Language Teaching* 44(2), 137-166.

Kakihana, S. 2024. Does the word-writing strategy in vocabulary learning increase learning outcomes or mislead learners into overestimation? *International Journal of Applied Linguistics* 34(3), 936-951.

Kapnoula, E. C. and A. G. Samuel. 2023. Wait long and prosper! delaying production alleviates its detrimental effect on word learning. *Language, Cognition and Neuroscience* 38(5), 724-744.

Laufer, B. and Z. Goldstein. 2004. Testing vocabulary knowledge: Size, strength, and computer adaptiveness. *Language Learning* 54(3), 399-436.

Lemhöfer, K. and M. Broersma. 2012. Introducing LexTALE: A quick and valid lexical test for advanced learners of English. *Behavior Research Methods* 44, 325-343.

Lee, E.-j. 2009. Vocabulary learning strategies for Korean EFL learners. *Studies in English Language & Literature* 35(1), 323-341.

Li, S. 2022. Working Memory and Second Language Learning: A Critical and Synthetic Review. In M. A. Bowles, T. Ionin and A. M. Perpiñán, eds., *The Routledge Handbook of Second Language Acquisition and Psycholinguistics* (pp. 348-360). Routledge.

Milton, J. and N. Hopkins. 2006. Comparing phonological and orthographic vocabulary size: Do vocabulary tests underestimate the knowledge of some learners? *Canadian Modern Language Review* 63(1), 127-147.

Perrone-Bertolotti, M., J. Kujala, J. R. Vidal, C. M. Hamame, T. Ossandon, O. Bertrand, L. Minotti, P. Kahane, K. Jerbi, and J.-P. Lachaux. 2012. How silent is silent reading? Intracerebral evidence for top-down activation of temporal voice areas during reading. *Journal of Neuroscience* 32(49), 17554-17562.

Peters, E. and S. Webb. 2018. Incidental vocabulary acquisition through viewing L2 television and factors that affect learning. *Studies in Second Language Acquisition* 40(3), 551-577.

R Core Team. 2024. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Ruiz, S., P. Rebuschat, and D. Meurers. 2021. The effects of working memory and declarative memory on instructed second language vocabulary learning: Insights from intelligent CALL. *Language Teaching Research* 25(4), 510-39.

© 2025 KASELL All rights reserved

Schmitt, N. 2008. Instructed second language vocabulary learning. *Language Teaching Research* 12(3), 329-363.
Vilhauer, R. P. 2017. Characteristics of inner reading voices. *Scandinavian Journal of Psychology* 58(4), 269-274.
Webb, S. 2008. The effects of context on incidental vocabulary learning. *Reading in a Foreign Language* 20(2), 232-245.
Webb, S. and A. Piasecki. 2018. Re-examining the effects of word writing on vocabulary learning. *ITL-International Journal of Applied Linguistics* 169(1), 72-94.
Wickham, H. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer.
Yoo, S. and H. Yoon. 2019. Effects of word-writing in contextual word learning on Korean high school students' L2 vocabulary learning. *The Journal of English Linguistics* 19(4), 713-735.

Examples in: English
Applicable Languages: English
Applicable Level: Tertiary

Appendix A

Table A1. Target Words and Nonwords Used in the Form Recognition Task

No.	Word	Meaning	Nonword
1	ardor	열정, 정열	ardoe
2	nimbus	비구름	nimbon
3	ladle	국자	ladre
4	offal	잡육, 찌꺼기 고기	offag
5	wraith	유령	wraithor
6	heretic	이단자	heretol
7	nostrum	엉터리 약	nostram
8	zephyr	산들바람, 미풍	zephyrn
9	bridle	굴레, 재갈	bridlow
10	quibble	트집, 불만	quibblet
11	waif	방랑자	weaf
12	mirth	웃음소리	mirthen
13	zealot	열성분자	zealor
14	paucity	소량, 부족	paucid
15	hubris	자만심	hubrix
16	proclivity	(흔히 좋지 못한) 성향	proclivin

© 2025 KASELL All rights reserved

Appendix B

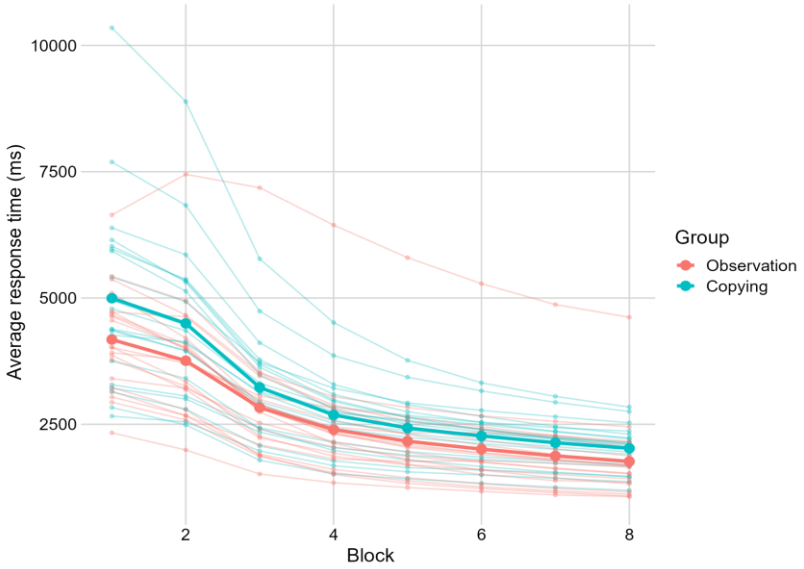

Figure B1. Mean Response Times per Block by Participant in the Meaning Recognition Task

Figure B1 presents the average response times of individual participants across blocks in the meaning recognition task administered after the target word learning phase. The bold lines represent the group-level means. Response times generally decreased over successive blocks, suggesting improved familiarity with the target word meanings.

Table B1. Accuracy Rates in the Meaning Recognition Task During Learning by Group

Group	Average accuracy (%)
OBSERVATION	99.61
COPYING	99.65

Note: Accuracy scores represent the percentage of correct responses in the meaning recognition task by training condition.

Table B1 summarizes the group-level accuracy rates during the learning phase. Both groups achieved a near-perfect average accuracy of 99.6%, confirming that participants successfully acquired the meanings of the target words.

© 2025 KASELL All rights reserved