The Reading Matrix: An International Online Journal
Volume 24, Number 2, September 2024

Word by Word: Investigating L2 Vocabulary Acquisition Through Extensive Reading

Brian Strong
Carleton University

ABSTRACT

While previous research has provided insights into vocabulary learning through extensive reading, the differential effects of word frequency and word class on active form and passive meaning word recognition remain less understood. By evaluating learners' post-test performance in active form recognition and passive meaning recognition, this study offers a comprehensive view of how these factors influence incidental vocabulary acquisition. The findings reveal a complex picture: while verbs showed a more significant improvement in active form recognition with increased exposure, underscoring the importance of repeated encounters, nouns were more effectively acquired in passive meaning recognition, suggesting differing cognitive processes involved in learning nouns versus verbs. Additionally, the study highlights that word frequency had a varied impact on learning outcomes, with multiple exposures facilitating a deeper understanding of verb forms, whereas noun learning showed less sensitivity to frequency. These results challenge the assumption that pleasure reading uniformly enhances all facets of word knowledge and suggest a complex interplay between word frequency, word class, and type of vocabulary recognition. The implications of these findings underscore the need for tailored strategies in L2 vocabulary instruction, advocating for a diversified approach to reading materials selection and emphasizing the importance of exposure diversity for effective vocabulary acquisition.

INTRODUCTION

The process of learning a second language (L2) hinges significantly on the acquisition of a robust vocabulary (Nation, 2022; Schmitt, 2009). Traditional methods of vocabulary instruction — such as memorization of bilingual lists, frequent dictionary referencing, and systematic vocabulary drills — have long been staples in L2 vocabulary learning (Nation, 2022). Despite their widespread use, these strategies are increasingly recognized as insufficient for fostering substantial vocabulary growth in L2 learners. They often fall short because they lack the rich, contextualized learning environments that mirror the natural language acquisition process, which is essential for deep and lasting language comprehension (Day & Bamford, 1998).

In response to these limitations, extensive reading has been posited as a powerful supplement to conventional vocabulary teaching methodologies (Horst, 2005). Unlike direct instruction, extensive reading encourages learners to absorb new vocabulary incidentally — a natural byproduct of immersive and meaningful engagement with language (Swanborn & De Glopper, 1999). When learners read widely for comprehension and pleasure, they encounter new

words in context, leading to vocabulary acquisition. The extensive reading approach involves learners choosing their reading material, thereby aligning with their interests and promoting motivation. This method not only aims to improve reading fluency and foster a lifelong reading habit but also inadvertently supports the acquisition of vocabulary.

This study seeks to examine the role of extensive reading in the incidental acquisition of vocabulary in an L2 context. More specifically, it investigates the impact that silent extensive reading has on the acquisition of active recognition of form and passive recognition of meaning of words, considering a range of word classes and frequencies. The research aims to provide empirical evidence supporting extensive reading as a viable method for vocabulary acquisition in L2 learning. Findings from this study could offer actionable insights for teaching practices, potentially enhancing the effectiveness of language education for both students and teachers.

LITERATURE REVIEW

Extensive reading is widely recognized in L2 acquisition as an approach that encourages learners to read large volumes of text for pleasure (Day & Bamford, 2004). This pedagogical strategy is grounded in the belief that reading broadly not only fosters general comprehension but also serves as a vehicle for the natural development of vocabulary, reading fluency, and intrinsic motivation (Iwata, 2022). Central to this approach is the concept of incidental vocabulary acquisition — whereby learners are thought to expand their lexical repertoire through exposure to new words within reading materials, absent of explicit vocabulary instruction.

The advocacy for extensive reading is predicated on its potential for incidental vocabulary gain, yet empirical studies present a complex tapestry of findings that often do not converge on a consensus regarding its effectiveness. It is within this context that factors such as the frequency of word exposure and the grammatical classification of words — nouns, verbs, adjectives, etc. — are postulated to play significant roles in the acquisition process (Schmitt, 2009).

Frequency effects on vocabulary learning

The relationship between word exposure frequency and vocabulary acquisition in L2 learners has garnered significant attention in the field of applied linguistics (e.g., Pellicer-Sandez & Schmitt, 2010). A strand of research suggests that vocabulary gains can be realized after minimal exposure to new words within reading materials. For instance, Rott's (1999) study provided early evidence of such effects among L2 learners. In this investigation, 95 German participants, enrolled in an intensive English program and categorized at an advanced beginner level, were exposed to 12 target words - an equal mix of nouns and verbs pertinent to daily life - distributed across six short narrative paragraphs. These words were assessed for both receptive and productive knowledge through a multiple-choice test and a fill-in exercise administered one week after the second exposure to each target word. Notably, the results indicated that significant gains in both types of vocabulary knowledge were achievable after just two exposures, underscoring the potential of incidental learning through reading and challenging the notion that extensive repetition is always necessary for vocabulary acquisition.

In a similar vein, Vidal's (2011) study extended the investigation of frequency effects, examining the vocabulary learning outcomes of 60 Spanish university students who engaged with a short story encompassing 10 target words drawn from the 2000-word frequency band.

Participants were divided into two groups, one encountering the story repeatedly over two weeks, and the other only once. The group with repetitive exposure demonstrated more robust vocabulary gains across measures of cued recall, recognition, and word familiarity, reinforcing the proposition that a moderate number of encounters (as few as 10 in this study) can facilitate L2 vocabulary acquisition.

Conversely, research by Waring and Takaki (2003) highlighted the necessity of more substantial repetition for word learning. Their study followed 39 Japanese university students who read a graded reader containing 25 low-frequency words over a period of 4-5 weeks. Follow-up tests showed that while form recognition reached 14% after two readings, approximately 7-8 exposures were required to achieve 50% recognition, and even after 18 exposures, form-meaning retention was less than 15%. This suggests a more complex interaction between frequency of encounters and vocabulary retention, indicating that numerous repetitions may be required to cement learning (Fujii, 2022).

Further complicating the understanding of frequency effects, Van Zeeland and Schmitt's (2013) study with 60 Dutch high school students found that vocabulary learning plateaued beyond a certain threshold of exposure. After reading fictional texts with 60 target words, the data revealed vocabulary improvements after 5-7 exposures, but additional encounters did not yield proportional gains, pointing to a potential saturation effect in vocabulary learning.

Additionally, studies such as those by Webb (2007) and Pigada and Schmitt (2006) have presented mixed results. Webb's investigation into the effects of varied versus single context exposures among 50 Japanese EFL learners suggested that learning is enhanced through repetition within multiple contexts, a finding that stands in contrast to linear models of vocabulary gains through increased frequency. Pigada and Schmitt's case study of a French learner highlighted a negligible correlation between frequency and acquisition across spelling, meaning, and grammar, with notable gains only manifesting when words were encountered 20 or more times.

In summary, the literature presents a complex picture of frequency effects on L2 vocabulary acquisition, with studies reporting a range of outcomes from significant gains after a few exposures for notable learning to occur. This variability underscores the intricate nature of vocabulary learning and signals the necessity for further research to dissect the contributory factors influencing frequency effects in extensive reading contexts.

Word class effects on vocabulary learning

The question of whether certain word classes (e.g., nouns, verbs, adjectives, etc.) are more readily acquired than others during vocabulary learning has been examined in multiple studies with mixed outcomes. Kweon and Kim (2008) investigated the incidental acquisition of 367 words — comprising nouns, verbs, and adjectives — by Korean high school EFL students. They employed a self-report vocabulary checklist to gauge knowledge gains post-exposure to the target words in various texts. The findings indicated a more robust retention of nouns in comparison to verbs and adjectives. However, the reliance on self-reporting as a measure introduced a degree of subjectivity, potentially skewing the accuracy of the reported gains. This study tentatively suggests a hierarchy in the ease of acquisition across word classes, with nouns being seemingly more accessible than other parts of speech, albeit with caution due to the methodological constraints of the measurement instrument.

Complementing this, Türk and Erçetin (2014) explored the incidental learning of 24 nouns and 24 verbs presented in a graded reader to Turkish EFL university students. These students, at a

pre-intermediate proficiency level, were assessed on both productive and receptive knowledge three weeks after reading the material containing these targeted lexical items. Their performance on these tests revealed a significantly higher acquisition rate for nouns compared to verbs. Nonetheless, the scope of lexical items and the singular text type used in this study limit the generalizability of these findings, suggesting that a broader array of words and text types might be necessary to fully understand the influence of word class on vocabulary learning.

In a similar investigation, Pellicer-Sánchez (2016) differentiated immediate recall abilities for nouns, verbs, adjectives, and adverbs among EFL university students with varying native languages and proficiency levels. This study reported the highest success rate for noun acquisition, followed by verbs, adverbs, and adjectives. However, it only assessed immediate recall and not long-term retention, thus limiting the conclusions that can be drawn about the enduring knowledge of these words (Bilici & Subasi, 2022).

In contrast, some research has found no distinct word class effects on vocabulary acquisition. Schoonen and Verhallen (2008) conducted a meta-analysis of vocabulary outcomes from multiple L2 reading studies and reported consistent vocabulary gains across different word classes. These results imply that factors beyond word class might exert a more substantial influence on vocabulary acquisition. Yet, it is important to note that the meta-analysis did not delve deeply into the potential interactions between word class and learner-specific variables, an area ripe for future inquiry.

Peters (2019) similarly reported no significant difference in the recall of concrete nouns and verbs by intermediate L2 English learners, indicating that word class alone may not be a determining factor in incidental vocabulary learning from reading. This study calls for further research with controlled designs that account for individual learner differences and utilize robust measures to clarify the effects of word class.

Selecting target vocabulary

In the domain of language learning, particularly in the context of L2 acquisition through reading, researchers have adopted various methodologies to explore the dynamics of incidental vocabulary learning. A pivotal aspect of such research involves the selection of target words within reading materials. This selection is crucial for isolating and measuring vocabulary acquisition without the interference of learners' pre-existing knowledge.

A novel approach to this challenge involves the use of pseudowords or lexical items that do not belong to the target language or the L2 learners' native language. An illustrative case is the study by Pitts et al. (1989), which incorporated nadsat words — fictional slang from the book A Clockwork Orange — as target words. These words, being absent from the learners' first languages, minimized the influence of language transfer on vocabulary learning. The use of such pseudowords ensures that any observed improvement in vocabulary knowledge stems directly from the reading activity, thereby providing a clearer picture of incidental vocabulary acquisition.

Waring and Takaki (2003) took a slightly different route by opting for substitute words that mimic plausible English vocabulary while adhering to conventional spelling rules. This choice was motivated by the belief that substitute words, by virtue of their resemblance to real English terms, would facilitate smoother reading and pronunciation for L2 learners. Unlike nonsense words, which might present pronunciation challenges or lack contextual cues for meaning deduction, substitute words blend into the reading flow without causing significant disruptions. Despite concerns that learners familiar with actual English words might find the substitutes challenging,

evidence suggested that comprehension was largely unaffected. Learners were able to deduce meanings for words like "yoot" (yes) and "molden" (dead), occasionally aided by their knowledge of the real English equivalents.

The strategic use of nonwords and substitute words in L2 vocabulary acquisition research offers key advantages. Primarily, it eliminates the variable of prior knowledge, ensuring that any vocabulary gains are attributable to the reading intervention. This approach also mitigates the effects of linguistic similarities between the learners' native language and the L2, thus offering a purer assessment of learning outcomes. Additionally, it facilitates comparative studies across different proficiency levels and linguistic backgrounds.

However, the use of artificial lexemes comes with its own set of limitations. The ecological validity of studies employing such words might be compromised, as these words do not reflect the authentic linguistic environment encountered by L2 learners. This raises questions about the applicability of research findings to real-world learning scenarios. Moreover, learners' motivation and engagement might suffer upon realizing they are learning words that lack practical utility. The artificial nature of these words could also interfere with the natural guessing processes involved in reading, potentially leading to confusion and inaccurate assessments of vocabulary acquisition.

In contrast to the aforementioned methodologies, Kweon and Kim (2008) focused their research on the acquisition of authentic words from unsimplified texts. This approach aimed to capture the natural process of vocabulary learning during extensive reading activities. By concentrating on real words that learners encountered in meaningful texts, the study offered insights into how vocabulary is naturally acquired and retained, presenting a more authentic perspective on L2 learning in real-life contexts.

In sum, the exploration of word frequency, word class, and incidental vocabulary acquisition within the context of second language (L2) learning through extensive reading reveals a complex landscape. Findings suggest that the effectiveness of vocabulary acquisition is influenced by the frequency of word exposure and the grammatical classification of words, highlighting a nuanced relationship that does not lend itself to simple generalizations. While certain studies advocate for the potential of significant vocabulary gains with minimal exposure, others underscore the importance of repeated encounters to solidify learning. Additionally, the acquisition ease across different word classes presents a varied picture; nouns may be acquired more readily than verbs or adjectives, though the evidence remains mixed and suggests the interplay of multiple factors. Methodological approaches, particularly the selection of target vocabulary and the use of pseudowords, are pivotal in navigating these complexities. This synthesis of empirical research not only identifies existing knowledge gaps but also calls for advanced investigative efforts to unravel the intricate dynamics of vocabulary learning through extensive reading. Such efforts aim to refine L2 pedagogical strategies and materials, optimizing the potential for vocabulary acquisition.

The Present Study

The literature review has highlighted a landscape of mixed findings concerning the impact of extensive reading on incidental vocabulary acquisition in a second language (L2) context. Although studies such as those by Rott (1999) and Vidal (2011) have suggested that extensive reading can lead to vocabulary gains, other research underscores the influence of various factors that complicate this process (Waring & Takaki, 2003; Pigada & Schmitt, 2006). Notably, there is no consensus on the optimal frequency of word exposure for maximum vocabulary retention, nor

on the differential effects of word classes. Additionally, the role of individual learner differences remains to be fully understood.

Given these ambiguities, this study is designed to probe the intricacies of L2 vocabulary development through extensive reading. Its focus is twofold: to dissect the effects of extensive reading on both active and passive word knowledge and to compare how word classes are acquired. It also seeks to refine our understanding of word frequency's role, taking into account the learners' proficiency levels.

The aim of this research is to bridge identified gaps by examining how extensive reading affects the incidental acquisition of both active and passive L2 vocabulary across various word classes and frequencies. The study will assess learning outcomes through recognition tests following the silent reading of graded texts over an extended timeframe. A detailed analysis of acquisition rates for nouns versus verbs at varying exposure levels will shed light on the interplay between reading habits, word class, word frequency, and lexical development. The outcomes will offer valuable insights for employing extensive reading in L2 instruction and contribute to designing more effective vocabulary learning strategies.

This investigation is guided by the following research questions:

- RQ1: What is the impact of extensive reading on the incidental acquisition of L2 vocabulary as measured by active form and passive meaning recognition tests?
- RQ2: How does extensive reading influence the incidental acquisition of words across different word classes, particularly nouns and verbs, in L2 contexts?
- RQ3: In what way does word frequency affect the incidental acquisition of L2 vocabulary within the framework of extensive reading?

Correspondingly, the following hypotheses were fomulated:

Hypothesis 1: Extensive reading will facilitate the incidental acquisition of new L2 words, with more substantial gains in passive recognition of meaning as compared to active recognition of form. This is expected because passive recognition more closely reflects the natural process of meaning comprehension during reading.

Hypothesis 2: Incidental L2 vocabulary learning from extensive reading will differ between word classes, with nouns likely being acquired more efficiently than verbs. This prediction is informed by previous studies which suggest that cognitive and contextual variables may influence the acquisition rates differently across word classes.

Hypothesis 3: Higher word frequency in reading materials will correlate with improved incidental vocabulary acquisition in L2. The expectation is that with increased exposure to words in diverse contexts, learners will strengthen their memory representations and engage in deeper processing, enhancing their retention and recall.

METHODOLOGY

Participants

This study involved the participation of 53 second-year university science majors who were learning English as a Foreign Language (EFL). The participants had an average age of 18.9 years, with 22 identifying as female. They were enrolled in a compulsory English language program at a private university in Japan, attending three one-hour English classes per week. The Vocabulary Levels Test (Schmitt et al., 2001) was administered to assess the participants' vocabulary knowledge, which indicated they had mastery of approximately 98%, 95%, and 89% of words

within the first, second, and third one-thousand frequency bands, respectively. Notably, the participants reported low motivation for learning English.

Prior to conducting this study, an a priori power analysis was performed using G*Power (Faul et al., 2009). The objective was to determine the minimum sample size required to detect medium effect sizes in a within-subjects ANOVA, with an alpha significance criterion set at .05 and a target power level of 80 percent. The analysis indicated that at least 34 participants would be necessary for robust hypothesis testing. The sample size of this study exceeded this threshold.

Materials

The materials for this investigation comprised two graded readers, selected for their suitability for intermediate-level EFL learners. The first text, "I, Robot," is a level 5 graded reader containing 1400 basic words, as categorized by MacMillan Readers. The second, "The Time Machine," encompasses approximately 1700 headwords and is classified as a level 4 graded reader within the Pearson English Readers' framework. Combined, these texts present learners with approximately 45,138 running words, offering a substantial corpus for extensive reading.

To evaluate the participants' pre-existing familiarity with the books' vocabulary, they were asked to read two pages from each book and note any unfamiliar words. This task was designed to estimate their baseline familiarity with the texts' vocabulary. Criteria for participation included a familiarity threshold with 98% of the running words in these samples. Following this preliminary assessment, a survey was administered to determine the participants' subjective enjoyment of the texts, which confirmed that both "I, Robot" and "The Time Machine" were received favorably by the participants.

Further analysis of the texts was facilitated by the use of Wordsmith Tools (version 6), which enabled the creation of a specialized corpus. Each word in the corpus was annotated for frequency and word class, from which 138 target words were selected. These target words spanned across five distinct frequency bands, ranging from words that occur only once to those that appear over sixteen times within the combined texts. The frequency bands and corresponding numbers of nouns and verbs within each are detailed in Table 1.

The selection process for these target words was guided by the discernment of an experienced English language teacher, who predicted the words that participants were least likely to be familiar with during reading. The teacher's intuition served as the sole criterion for this selection, aiming to mimic the authentic challenge learners face when encountering potentially unfamiliar vocabulary within a reading context.

Table 1. Word Frequency and Word Class of Target Items

Frequency band	Words per Band	Word Frequency
1	42 (25 nouns, 17 verbs)	1 occurrence
2	27 (13 nouns, 14 verbs)	2-4 occurrences
3	21 (13 nouns, 8 verbs)	5-8 occurrences
4	22 (14 nouns, 8 verbs)	9-15 occurrences
5	26 (17 nouns, 9 verbs)	16+ occurrences

Pre- and Post-tests

The study deployed two specific types of tests to assess the effects of extensive reading on incidental vocabulary acquisition in an L2: passive recognition of meaning and active recognition of form. Following Laufer and Goldstein's (2004) hierarchy of vocabulary knowledge, passive recognition of meaning assesses the ability to understand the meaning of an L2 word in a receptive context, such as reading or listening. This level of lexical understanding is crucial for language comprehension and often serves as the initial gauge of L2 proficiency. Conversely, active recognition of form requires a deeper level of engagement with the language, demanding learners to recognize the form of an L2 word when presented with its L1 translation equivalent.

For the active recognition of form test, participants received a cue in their native language (Japanese) and chose the corresponding English word from a set of options. To minimize random guessing, an "l don't know" option was included. For example, given the cue 炎 (flame), participants selected from options such as: a) apple, b) flame, c) string, d) lawn, e) sky, f) I don't know. The distractors were designed to be semantically and phonologically dissimilar to the target word, ensuring a consistent level of challenge while controlling for word frequency.

In the passive recognition of meaning test, participants were given an English word (e.g., flame) and asked to select the correct Japanese translation from multiple choices, such as: a) リンゴ (apple), b) 炎 (flame), c) ひも (string), d) 芝生 (lawn), e) 空 (sky), and f) 分かりません (I don't know).

The tests were administered at two intervals: three weeks before the start of the reading treatment (pre-test) to establish a baseline, and one week after its conclusion (post-test) to measure the vocabulary gains. The pre-test was scheduled well in advance to reduce the likelihood of a testing effect. The post-test's timing, while dictated by the class schedule, also offered an appropriate balance between testing immediately after learning and allowing for a short consolidation period.

In both pre- and post-test phases, the active recognition of form test was administered before the passive recognition of meaning test. The pre-test was conducted without offering corrective feedback to prevent any learning effects that might impact the study's integrity. The post-test, however, included feedback to assist in the acquisition of unfamiliar words from the graded readers, in line with practices that enhance ecological validity.

An additional layer of rigor was introduced in the post-test through the randomization of test items for controlling for potential order effects — an approach aligns with best practices for test administration in second language acquisition research. Furthermore, to ensure that participants' performance reflected their true vocabulary knowledge, no time limit was imposed on either test.

Procedure

This study's methodology was strategically divided into four distinct phases: the Vocabulary Levels Test (VLT), the pretest, the reading treatment, and the post-test. Each phase was essential for evaluating the efficacy of reading on incidental vocabulary acquisition.

The initial phase commenced with the administration of the VLT (Schmitt, Schmitt, & Clapham, 2001) four weeks before the reading treatment. This test measures receptive vocabulary across five graded frequency levels. Each level presented 30 target words, and participants

indicated their understanding by selecting the correct meanings from a set of six options. Performance was quantified as a percentage of correct answers for each frequency level. One week after the VLT, the pretest was administered. The pretest aimed to map out baseline knowledge of the target vocabulary. This was achieved through a combination of passive and active recognition tests.

Subsequently, the reading treatment was integrated into the beginning of regular English language class sessions over a three-week period. Each session incorporated a 45-minute segment dedicated to silent reading of graded readers, with the aim of boosting reading fluency. Participants were instructed to sequentially read two books, with the order randomly assigned to each student to prevent order bias. After the designated reading time, the materials were stored away, and the usual class lesson was resumed. The cumulative duration of this treatment amounted to approximately 6.8 hours.

The final phase, conducted one week after the reading treatment, involved an unexpected post-test to assess vocabulary retention. This delay not only allowed for the consolidation of newly acquired vocabulary—thus supporting the transition from short-term to long-term memory. By employing a surprise post-test, the study minimized potential biases from explicit memorization tactics, focusing instead on the natural, incidental vocabulary acquisition from the reading treatment.

Scoring and data analysis

In this research, the scoring of test responses was binary, a common approach for measuring accuracy in recognition tests. Each correct response by a participant was assigned a score of ‘1’, whereas incorrect responses were scored ‘0’. This dichotomous scoring system facilitated a clear, objective measure of participant comprehension and retrieval.

To ensure the integrity of the data, a rigorous treatment was applied to account for the participants’ prior knowledge. In alignment with the methodology of Yamagata et al. (2022), responses that were correctly answered in the pretest phase were subsequently treated as missing data. This decision was crucial to the study's validity, as it isolated the effects of the reading treatment on vocabulary acquisition by removing the potential confounding variable of pre-existing knowledge.

The core of the data analysis was conducted using a Generalized Linear Mixed Model (GLMM) with the GAMLj module in Jamovi, a robust statistical package. The GLMM was specified with fixed effects for ‘Word Frequency’, which was stratified into 5 distinct levels, and ‘Word Class’, categorized into 2 levels. The interaction between these two variables was also included in the model to understand their combined effect on active form and passive meaning recognition. Additionally, random intercepts for both subjects and targets were incorporated to account for inter-individual differences and item-specific variance, respectively.

The model’s complexity was thoughtfully considered. Although the inclusion of random slopes for the fixed effects was contemplated, they were ultimately excluded after model fit comparisons using the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) did not justify their inclusion. Models with and without random slopes were rigorously tested, and those with only random intercepts were deemed to offer the best fit based on these statistical criteria.

For the interpretation of the GLMM results, the odds ratios obtained were translated into effect sizes using Cohen’s *d*. The conversion employed Chinn’s (2000) formula: ln(odds

ratio)/1.81. The effect sizes were then categorized according to the benchmarks suggested by Plonsky and Oswald (2014), which designate an effect size of 0.4 as small, 0.7 as medium, and 1.0 or higher as large. These benchmarks offer a standardized method for evaluating the practical significance of the results.

During the analysis phase, the VLT scores, which served as a baseline measure of vocabulary, were considered as a potential covariate. However, model fit comparisons using AIC and BIC concluded that the inclusion of VLT scores did not significantly enhance the model's explanatory power, leading to their exclusion from the final analysis.

RESULTS

This study aimed to uncover patterns in incidental L2 vocabulary acquisition through reading by analyzing data from both active form and passive meaning recognition tests. The data revealed interesting trends in how participants learned words, particularly concerning word class and the frequency of encounters with these words in texts.

The findings, as summarized in Table 2, suggest a difference in learning outcomes based on word class. On average, verbs were learned more effectively than nouns. More specifically, the correct recognition rate for verbs stood at 26% (*SD* = 44%), compared to a 21% (*SD* = 41%) rate for nouns. The frequency with which words were encountered also played a significant role in vocabulary acquisition. Participants showed an average accuracy of 26% (*SD* = 44%) when encountering words only once. This accuracy improved to 43% (*SD* = 50%) with 2-4 encounters. Surprisingly, further encounters (5-8 times) resulted in a decreased average accuracy of 20% (*SD* = 40%), and this decline continued with 9-15 encounters, dropping the average accuracy to 14% (*SD* = 35%). Words encountered 16 or more times saw a similar average accuracy of 13% (*SD* = 34%), suggesting that increased exposure beyond a certain frequency may not continue to benefit learning in the same way.

Table 2. The Percentage of Correct Responses on the Active Form and Passive Meaning Recognition Post-tests

	Encounters	**Passive Recognition**		**Active Recognition**	
		M	*SD*	*M*	*SD*
Nouns	1	56%	50%	50%	50%
	2-4	52%	50%	56%	49%
	5-8	60%	49%	05%	21%
	9-15	44%	50%	02%	15%
	16+	58%	50%	01%	11%
Verbs	1	36%	48%	04%	20%
	2-4	41%	49%	33%	47%
	5-8	47%	50%	71%	46%
	9-15	41%	50%	70%	46%
	16+	47%	50%	45%	50%

Active Recognition of Form Test Results

The results of the GLMM, as illustrated in Figure 1, indicate that Word Class significantly impacted post-test performance in the active form recognition test (χ^2 = 37.5, p < .001).

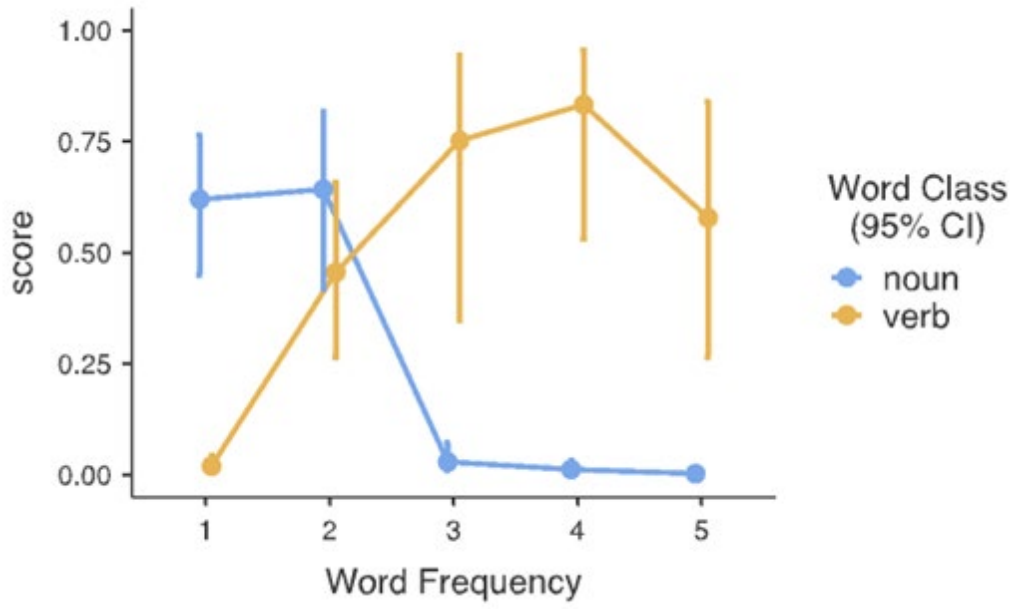
Specifically, Bonferroni post hoc comparisons revealed that verbs were learned significantly better than nouns, ($z = -6.13$, $p < .001$), suggesting that the likelihood of correctly recognizing the form of a verb was notably higher than that of a noun.

In the active recognition of form test, Word Frequency was also found to be a significant factor ($\chi^2 = 34.1$, $p < .001$). Bonferroni post hoc comparisons revealed that when comparing the frequency of word encounters from one to 2-4, there was a significant increase in the likelihood of correct recognition ($z = -4.607$, $p < .001$). However, a significant decrease in correct recognition was observed when comparing encounters of words 2-4 times to 9-25 times ($z = 2.39$, $p = 0.03$) and an even more substantial decrease when compare 2-4 encounters to sixteen or more encounters ($z = 5.23$, $p < .001$). These results suggest a complex relationship between exposure frequency and the active recognition of word forms, where both too few and too many exposures may be less effective than a moderate number of encounters.

The model also reported a statistically significant interaction between Word Frequency and Word Class ($\chi^2 = 145.0$, $p < .001$). Bonferroni post hoc analyses exploring the interaction between word class (nouns and verbs) and exposure frequency yielded several significant findings. This interaction was found to be significant for both nouns and verbs at various levels of exposure, with distinct patterns emerging for each word class.

For nouns, there was a significant decrease in active recognition of meaning performance when comparing the baseline of one exposure to the 5-8 exposure category ($z = 6.543$, $p < .001$). This trend continued, with substantial decreases also observed when comparing one exposure to 9-15 ($z = 7.78$, $p < .001$) and to 16+ exposures ($z = 8.66$, $p < .001$). Notably, the active form recognition scores peaked at the 2-4 exposure frequency before declining, suggesting an optimal frequency range for noun acquisition. In contrast, verbs showed an escalating pattern of active form recognition with increased exposure. A significant improvement was noted as frequency increased from the one exposure category to the 2-4 category ($z = -6.13$, $p < .001$).

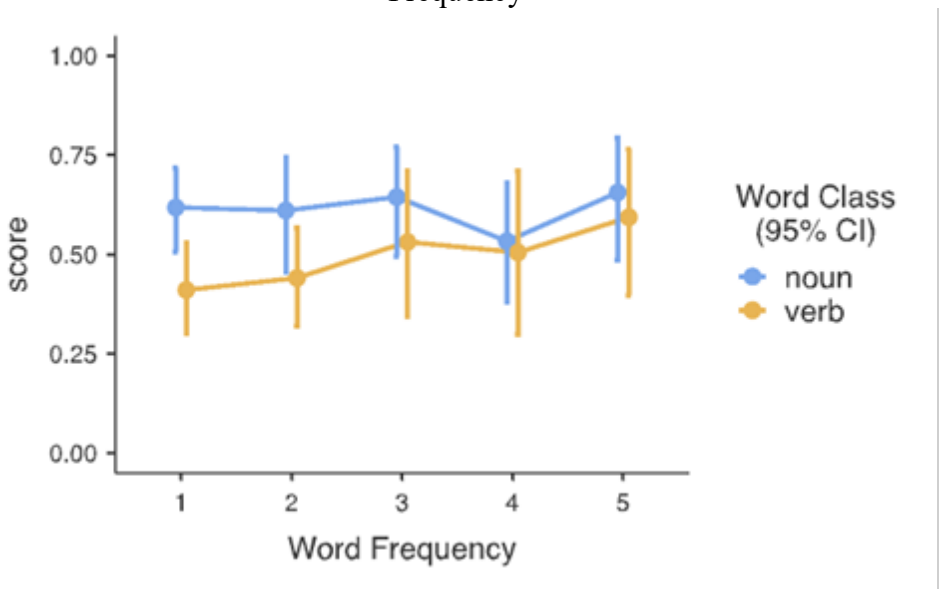
Figure 1. Results of the Active Form Recognition Test by Word Class and Word Frequency



Passive Recognition of Meaning Test Results

In examining the fixed effects within the GLMM, the omnibus test revealed a significant effect of Word Class on performance in the passive recognition of meaning test ($\chi^2 = 5.24$, $p = .02$), with nouns being more likely to be recognized than verbs. However, Word Frequency did not significantly affect the outcome ($\chi^2 = 4.07$, $p = .39$). Additionally, the interaction between word class and word frequency was not significant ($\chi^2 = 1.94$, $p = .75$).

Figure 2. Results of the Passive Recognition of Meaning Test by Word Class and Word Frequency



DISCUSSION

The present study contributes to our understanding of the dynamics of incidental vocabulary learning through sustained silent reading in an L2 context. Our exploration centered around three research questions, which examined the effects of pleasure reading on vocabulary acquisition, differences in retention across word classes, and the influence of encounter frequency on word retention.

The data from this study offers partial support for the first hypothesis, which posited that L2 learners would incidentally acquire unknown words through extensive reading, with a more pronounced effect on passive meaning recognition than active form recognition. The findings revealed that a single encounter with nouns led to a mean accuracy rate of 50% for active form recognition and 56% for passive meaning recognition, while verbs showed a mean accuracy rate of 4% for active form recognition and 36% for passive meaning recognition. This suggests that learners are somewhat more likely to recognize the meaning of words (passive meaning recognition) than to recall their form (active form recognition) after one exposure, which supports the natural process of meaning access during reading (Horst, Cobb, & Meara, 1998).

However, as learners encountered words more frequently, particularly verbs, there was a substantial increase in active form recognition, peaking at 71% accuracy after 5-8 encounters. This trend diverges from the initial hypothesis, as it indicates that repeated exposure enhances active form recognition significantly, particularly for verbs. Such findings echo the research by Webb (2007), which underlined the importance of repetition for word form retention.

The distinction in performance between active form and passive meaning recognition tests aligns with the differing cognitive processing demands associated with each task. Active form recognition, prompted by an L1 cue, may facilitate a more direct retrieval of the L2 word form, while passive meaning recognition requires learners to link an L2 word to its meaning from multiple choices, which can be more cognitively demanding (Laufer & Goldstein, 2004). This difference underscores the complex nature of L2 lexical recognition and suggests that the process is influenced by the cognitive demands of the task at hand.

Hypothesis 2, which suggested differential effects of extensive reading on nouns versus verbs, was partially supported. Active recognition tests showed better learning for nouns, which resonates with previous research suggesting that nouns are typically acquired more easily than verbs during extensive reading (L1 research: Tomasello, 2000; Waxman et al., 2013). However, in the passive recognition tests, an unexpected pattern emerged where verbs outperformed nouns. This might reflect the differential semantic richness between nouns and verbs, with verbs potentially offering more contextual cues that aid retention (Gentner, 2006).

In addressing Hypothesis 3, our investigation sheds light on the intricate relationship between word frequency and incidental vocabulary acquisition through extensive reading. The hypothesis, which anticipated a positive correlation between higher word frequency and enhanced vocabulary acquisition, was corroborated only to a certain extent. The empirical evidence revealed a subtle interaction wherein vocabulary acquisition initially benefits from increased word frequency, reaching an optimal threshold beyond which additional exposure does not yield proportional gains in learning. This discovery not only punctuates the critical role of word frequency in designing effective extensive reading interventions but also invites a reevaluation of how reading materials are curated to maximize vocabulary acquisition among L2 learners.

Pedagogical Implications

The findings of this study have several important implications for L2 vocabulary instruction. The results highlight the nuanced impact of word frequency and word class on vocabulary acquisition through extensive reading, suggesting that these factors should be carefully considered when designing language learning activities.

The positive effects of extensive reading on vocabulary acquisition emphasize the need for incorporating extensive reading programs into L2 curricula. Allowing students to choose reading materials that interest them can enhance their motivation and engagement. However, while extensive reading is beneficial, it is important for educators to recognize that it does not guarantee the acquisition of all encountered vocabulary. This study found that verbs showed significant improvement in active form recognition with increased exposure, whereas nouns were more effectively acquired in passive meaning recognition. Consequently, educators should ensure that reading materials include a balanced mix of word classes and frequencies to cater to the differing cognitive processes involved in learning various types of words.

Teachers should design activities that provide students with varied exposure to target words in different contexts. Reading circles, paired reading sessions, and interactive reading logs can help facilitate these exposures. Given the study's findings on the differential acquisition of nouns and verbs, teaching strategies should be tailored accordingly. For example, interactive storytelling and role-playing can be particularly effective for verb acquisition, while labeling and descriptive writing exercises can support the learning of nouns.

While incorporating extensive reading is crucial, educators must supplement it with other vocabulary learning strategies to ensure comprehensive vocabulary development. Explicit vocabulary instruction, focused practice, and opportunities for active use of new words can help solidify learning that may not occur solely through reading.

LIMITATIONS AND CONCLUSION

The current study presents insightful evidence on the positive impact of extensive reading on incidental second language (L2) vocabulary acquisition among non-English major university students in Japan. While bolstering the academic narrative on vocabulary learning in higher education, it is critical to discuss the limitations that frame the interpretation of the findings and propose directions for future research.

The study’s cohort, comprising second-year science majors from a single private university in Japan, fulfilled the a priori power analysis requirements. However, this specific demographic limits the generalizability of the results. Participants’ non-English majors status and their reported disinterest in English learning raise questions about the applicability of findings to other learner populations. Future investigations should aim to involve a broader demographic, including students from various academic disciplines and with varying levels of motivation towards English language acquisition, from multiple universities to enhance the external validity of the research.

The reliance on two specific graded readers based on participants’ preferences introduced a constraint in terms of material diversity. While this choice ensured engagement with the texts, it narrowed the scope of genres and themes explored, potentially influencing the transferability of the findings to other reading materials. To address this, future research should diversify the selection of texts, incorporating a range of genres and complexity levels to cater to different reader profiles and interests.

The methodology employed for selecting target vocabulary items was subjectively informed by an English language teacher's intuition. This approach, although expert-informed, is susceptible to bias. Future research methodologies could benefit from more objective, possibly corpus-based, selection strategies to enhance the validity of the vocabulary items chosen for study.

The study's assessment focused on active and passive recognition, omitting productive vocabulary usage in speaking and writing. This represents a notable gap, as productive skills are crucial in language acquisition. Subsequent studies should include assessments of productive vocabulary knowledge to offer a more comprehensive evaluation of vocabulary learning outcomes.

The study lacked a rigorous mechanism to verify the completion and comprehension of the reading materials, relying on participant self-reporting. This casts potential doubt on the reliability of the reported reading outcomes. To mitigate this, future research could integrate more robust verification methods, such as engaging discussions or written summaries, to ensure and validate reading completion and comprehension.

The estimation of reading rates was based on the maximum time available during class sessions, which may not reflect the actual time spent on reading. Accurate time tracking in future studies could offer a clearer picture of the time investment in reading and its correlation with vocabulary acquisition.

The multiple-choice format employed for assessing active recall may have inadvertently allowed for inflated scores due to guessing, despite measures to counteract this. The simplicity of the response options might not accurately capture the nuances of learners' knowledge. A reviewer's suggestion to use translation tests in subsequent research is acknowledged as a means to reduce the potential for test design artifacts and obtain a more precise measure of lexical knowledge.

Lastly, the absence of qualitative insights limits a deeper understanding of participants' engagement with the texts, motivations, and learning strategies. Future research could benefit from incorporating qualitative methods, such as interviews or think-aloud protocols, to gain a richer, more nuanced understanding of the processes underlying vocabulary acquisition.

Despite these limitations, the study makes a significant contribution to the understanding of L2 vocabulary learning within an academic context. The demonstrated benefits of extensive reading among non-English major undergraduates in Japan provide valuable insights, particularly in the realm of incidental vocabulary acquisition. The findings support the notion that extensive reading can be a powerful tool for vocabulary development in L2 learners, aligning with Brown et al.'s (2008) data, which emphasize the influence of testing formats on learning outcomes. As such, this research both underscores the importance of methodological considerations in vocabulary acquisition studies and serves as a foundational reference for future explorations into the efficacy of extensive reading programs. The limitations identified herein not only contextualize the current study's results but also chart a course for future inquiries to enrich our collective understanding of the dynamics of vocabulary learning in L2 education.

REFERENCES

Bamford, J., & Day, R. R. (Eds.). (2004). *Extensive reading activities for teaching language*. Cambridge University Press.

Bilici, Ö. U., & Sabaşi, G. (2022). EFL Learners' metacognitive awareness of global reading strategies and L2 reading comprehension skills: The efficacy of reciprocal teaching. *The Reading Matrix: An International Online Journal*, *22*(2), 170–185.

Brown, R., Waring, R., & Donkaewbua, S. (2008). Incidental vocabulary acquisition from reading, reading-while-listening, and listening to stories. *Reading in a Foreign Language*, *20*(2), 136–163.

Day, R. R., & Bamford, J. (1998). *Extensive reading in the second language classroom*. Cambridge University Press.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Fujii, K. (2022). The effects of a two-year-long extensive reading program on TOEIC Bridge IP scores. *The Reading Matrix: An International Online Journal*, *22*(2), 108–120.

Horst, M. (2005). Learning L2 vocabulary through extensive reading: A measurement study. *Canadian Modern Language Review*, *61*(3), 355–382.

Horst, M., Cobb, T., & Meara, P. (1998). Beyond a clockwork orange: Acquiring second language vocabulary through reading. *Reading in a Foreign Language*, *11*(2), 207–223.

Iwata, A. (2022). The effectiveness of extensive reading (ER) on the development of EFL learners' sight vocabulary size and reading fluency. *The Reading Matrix: An International Online Journal*, *22*(2), 74–91.

Kweon, S.O., & Kim, H.R. (2008). Beyond raw frequency: Incidental vocabulary acquisition in extensive reading. *Reading in a Foreign Language*, *20*(2), 191–215.

Laufer, B., & Goldstein, Z. (2004). Testing vocabulary knowledge: Size, strength, and computer adaptiveness: Language learning. *Language Learning*, *54*(3), 399–436. https://doi.org/10.1111/j.0023-8333.2004.00260.x

Nation, I. S. P. (2022). *Learning Vocabulary in Another Language* (3rd ed.). Cambridge University Press. https://doi.org/10.1017/9781009093873

Pellicer-Sanchez, A., & Schmitt, N. (2010). Incidental vocabulary acquisition from an authentic novel: Do things fall apart? *Reading in a Foreign Language*, *22*(1), 31–55.

Pigada, M., & Schmitt, N. (2006). Vocabulary acquisition from extensive reading: A case study. *Reading in a Foreign Language*, *18*(1), 1–28.

Pitts, M., White, H., & Krashen, S. (1989). Acquiring second language vocabulary through reading: A replication of the Clockwork Orange study using second language acquirers. *Reading in a Foreign Language*, *5*(2), 271–275.

Rott, S. (1999). The effect of exposure frequency on intermediate language learners' incidental vocabulary acquisition and retention through reading. *Studies in Second Language Acquisition*, *21*(4), 589–619. https://doi.org/10.1017/S0272263199004039

Schmitt, N. (Ed.). (2009). *Vocabulary: Description, acquisition and pedagogy*. Cambridge University. Press.

Swanborn, M. S. L., & De Glopper, K. (1999). Incidental word learning while reading: A meta-analysis. Review of Educational Research, *69*(3), 261–285. https://doi.org/10.3102/00346543069003261

Waring, R., & Takaki, M. (2003). At what rate do learners learn and retain new vocabulary from reading a graded reader? *Reading in a Foreign Language*, *15*(2), 130–163.

Waxman, S., Fu, X., Arunachalam, S., Leddon, E., Geraghty, K., & Song, H. (2013). Are nouns learned before verbs? Infants provide insight into a long-standing debate. *Child Development Perspectives*, *7*(3), 155–159. https://doi.org/10.1111/cdep.12032

Yamagata, S., Nakata, T., & Rogers, J. (2022). Effects of distributed practice on the acquisition of verb-noun collocations. *Studies in Second Language Acquisition*, 1–27. https://doi.org/10.1017/S0272263122000225

Brian Strong *is an Assistant Professor in the School of Linguistics and Language Studies at Carleton University, Canada. His interests are EFL reading, vocabulary acquisition and learning methods.*

Email: brian.strong@carleton.ca