The effect of extensive reading and paired-associate learning on long-term vocabulary retention: An event-related potential study

Eunjin Chun, Sungmook Choi*, Junsoo Kim
Teachers College 502-4, Kyungpook National University, 702-701, South Korea

HIGHLIGHTS
► This study reports the first ERP findings in the extensive reading literature.
► Extensive reading led to significant long-term vocabulary retention.
► Paired-associated learning was effective only in short-term vocabulary retention.
► These findings are supported by the N400 component of event-related potentials.
► These findings are also supported by behavioral measures.

ARTICLE INFO
Article history:
Received 19 April 2012
Received in revised form 12 May 2012
Accepted 28 May 2012

Keywords:
Extensive reading
Paired-associate learning
Vocabulary
Event-related potential
N400

ABSTRACT
We investigated the relative efficacy of extensive reading (ER) and paired-associate learning (PAL) in the ability of second language (L2) learners to retain new vocabulary words. To that end, we combined behavioral measures (i.e., vocabulary tests) and an event-related potential (ERP) investigation with a focus on the N400 ERP component to track short- and long-term vocabulary retention as a consequence of the two different approaches. Behavioral results indicated that both ER and PAL led to substantial short-term retention of the target words. In contrast, on a long-term basis, ER was more effective than PAL to a considerable degree as indicated by a large-size effect (d = 1.35). Evidence from the N400 effects (d = 1.70) observed in the parietal electrode group (P3, Pz, P4) provided further support for the superior effects of ER over PAL on long-term vocabulary retention. The converging evidence challenges the assumptions of some L2 researchers and makes a significant contribution to the literature of vocabulary acquisition, because it provides the first ERP evidence that ER is more conducive to long-term vocabulary retention than PAL.

© 2012 Elsevier Ireland Ltd. All rights reserved.

1. Introduction

There is a growing consensus among second language (L2) researchers that sufficient lexical knowledge is a critical component of L2 proficiency [7,10,22]. Coady and Huckin [3] claim that L2 learners should acquire a substantial vocabulary in order to attain competencies in language skills such as listening, speaking, reading, and writing. In a related vein, learners who do not have a large repertoire of vocabulary knowledge typically avoid reading and may struggle throughout their schooling [8].

A vast body of research has established that extensive reading (ER) is a highly effective means of vocabulary development, especially for first language (L1) learners [5,17,18]. As an illustration, Cunningham and Stanovich [5] found that the amount of time spent on reading is one of the most potent predictors of vocabulary size for learners in grades 4, 5, and 6. Nagy et al. [18] concluded that an average amount of reading probably accounts for a third of the vocabulary growth of a child, and that regular and extensive reading is a strong predictor of substantial and permanent vocabulary development.

Several L2 researchers also report evidence in support of ER [12,20]. For instance, Krashen [12] finds that vocabulary learning through meaning-focused reading significantly contributes to vocabulary building. Similarly, Nation and Wang [20] and Warling and Takaki [25] also find that the reading of graded readers is effective in vocabulary development.

Counter-arguments, however, were also put forward by other L2 researchers [2,9,14–16,19,23]. Several researchers reported that ER is not an efficient means of L2 vocabulary acquisition. For instance, in Zahar’s [27] study, grade seven L2 learners were asked to read a story entitled Golden Fleece and respond to a vocabulary test. The results indicated no significant effect of story reading on word
acquisition. Moreover, many L2 researchers found that decontextualized vocabulary learning methods such as mnemonic strategies and repeated rehearsals are far more effective than learning words from context [15,16,23]. For instance, Qian [23] compared the effects of decontextualized word lists and contextualized vocabulary on vocabulary acquisition. The researcher found that learners in a decontextualized condition learned more words than did their counterparts in a contextualized condition. More recently, Lindstromberg and Boers [16] showed that a mnemonic strategy using alliteration resulted in significant lexical gain. Laufer and Girsai [15] found that contrastive semantic analysis (e.g., comparing the degree of overlap between L1 and L2 words) yields a better outcome than meaning-oriented tasks for the short- and long-term learning of new words. In sum, the trends in L2 vocabulary acquisition have gravitated toward explicit attention to target words in a decontextualized manner and through frequent rehearsals [9,14].

Despite the work of these researchers to address the impact of the ER and decontextualized vocabulary learning, three questions still remain unanswered. First, little knowledge is shed on the relative efficacy of ER and decontextualized vocabulary learning in light of short- and long-term vocabulary retention. Second, although paired-associate learning (PAL), that is, learning new foreign words with their translations or synonyms, is one of the most popular methods in decontextualized vocabulary learning, no previous studies have compared the effects of PAL with that of ER on vocabulary acquisition. Research has consistently shown that PAL is an effective means of acquiring new words [19,20,26]. For instance, Webb [26] compared relative efficiency of receptive and productive PALs with Japanese L2 learners. The results showed that productive PAL was more effective than receptive PAL in the acquisition of new words. In fact, PAL represents the most prevailing strategy for the participants in this study (Korean students who learn English as a foreign language), whereas ER represents the least-favored vocabulary learning approach for the same participants. Finally, online evidence using event-related potentials (ERPs) has never been demonstrated previously to determine the effects of ER and PAL. To better understand the effects of ER and PAL, it is imperative to identify electrophysiological responses underlying the potentially different effects of ER and PAL.

2. Present study

Two aspects set this study apart from previous studies. First, unlike previous studies, the present study directly compared the relative efficacy of ER and PAL in the same study. Second, we measured online vocabulary processes by recording electrophysiological responses. Specifically, we performed a priming experiment (i.e., prime–target paradigm) with a specific focus on the N400 component of ERPs. The N400 is a negative brain potential which peaks approximately 400 ms after the onset of a stimulus (hence the term N400). It is correlated with how well a word (i.e., target) semantically fits in with previous stimuli (i.e., prime), such as a word, a sentence, discourse, or image. The poorer the fit, the larger the N400 effect; the better the fit, the more reduced the N400 effect [13]. In other words, when a word is congruent with its preceding stimuli (e.g., shark – fish), the N400 effects are reduced. Conversely, when a word is unrelated with its previous stimuli (e.g., shark – sky), the N400 effects are increased.

To illustrate, a recent study by Schmidt-Kassow et al. [24] focused on the N400 component to determine the effects of exercising on vocabulary learning. The findings indicated that participants who were physically active during vocabulary learning performed better in vocabulary tests and also showed a larger N400 effect over central and right hemispheric electrode sites in reaction to the prime–target mismatch condition (e.g., apple–dog). In our study, the presence of an N400 effect elicited by prime–target violations would reveal the relative effectiveness of ER and PAL in vocabulary learning. For instance, if ER were more effective than PAL, ER participants would show a larger N400 effect than PAL for mismatched prime–target pairs.

3. Methods

3.1. Participants and design

In the current study, twenty-six adult native Korean speakers volunteered to take part in a 9-week experiment (twelve subjects were excluded from data analysis because they did not complete all the tasks). We offered small gifts in appreciation of their participation instead of monetary compensation. In order to control for confounding variables, the participants were screened in terms of handedness, age, gender, language proficiency, and knowledge of target words. For instance, all participants were right-handed with no history of neurological or psychiatric disorders, and their eyesight was normal or corrected to normal vision. They all agreed to avoid any intentional English-related activities throughout the duration of experiment. The participants were pseudo-randomly assigned to one of two exposure conditions, ER and PAL. The two groups were matched for age (mean age = 25) and language proficiency (low intermediate level).

This study employed a between-subjects pre-test, immediate post-test, and delayed post-test design. Specifically, this study is composed of an exposure period (i.e., ER and PAL for five weeks, respectively) and three test phases: (i) pre-exposure phase, (ii) immediate post-exposure phase, and (iii) delayed post-exposure phase (four weeks after immediate post-exposure phase). At each phase, participants performed a vocabulary test and responded to prime–target stimuli while their EEG activities were recorded. The vocabulary test and stimuli were identical across the three test phases.

3.2. Procedures

The 9-week experiment was conducted in the following way. First, we measured participant knowledge of 80 target words (plus 70 dummy words) before exposure (i.e., ER and PAL) in order to reduce the variability due to the two groups’ knowledge of target English words. The result showed that the two groups did not differ significantly (t = -.658, p = .523), indicating that the two groups were homogenous in the knowledge of the target words.

Second, ER and PAL groups were exposed to the same target words according to their respective condition. Specifically, the ER group participants were asked to read a book per week (1 book × 5 weeks = 5 books). The selected books were five graded readers at level three from Penguin Young Readers (three fiction and two nonfiction books) – Sherlock Holmes and other stories, Sense and Sensibility, Titanic, A Young King and other stories, and The British Life. Participants were allowed to select the order to read the books. They were also allowed to use a dictionary when they encountered unknown words. However, they were not told which words they would be tested on after five weeks of reading. To ensure that they read all of the books, we examined their comprehension of the main ideas after each reading.

In the PAL group, participants were asked to memorize thirty word pairs in two languages each week (such as statue – 동상). All told, they were requested to learn 150 words by heart. As with ER group participants, the PAL group participants were also not told which words they would be tested on. In order to ensure that they memorized thirty target words per week, we followed up on their learning.
Third, after a 5-week exposure to target words, the ER and PAL groups performed an immediate post-exposure vocabulary test and responded to the prime-target stimuli for EEG recording. In the vocabulary test, participants were asked to write down the Korean translation of the word. This task was immediately followed by an EEG recording. Finally, in the 9th week, the two groups took a delayed post-exposure vocabulary test and again responded to the prime-target stimuli for final EEG recording.

3.3. EEG recording and stimuli

During the recording, participants were seated comfortably in a sound-attenuated, electrically shielded room. Except for rest periods, they were requested to control blinking, swallowing, and other muscle movements to ensure the quality of the EEG data. Based on the 10/20 international system [11], EEG activities were continuously recorded from 15 scalp locations (O1, O2, Fz, F3, F4, Cz, C3, C4, Pz, P3, P4, T3, T4, T5, and T6) by means of Ag/AgCl electrode caps. The electrode impedance threshold value was kept below 5 kΩ using a Glass ohmmeter. Eye movements were monitored via an electrode placed at the outer canthus of the left eye, while vertical eye movements and blinks were recorded by means of an electrode beneath the right eye. The electrode Fz on the cap served as grounding, while a reference electrode was placed on the right ear lobe. Signals were amplified 50,000 times with MP150 (BIOPAC System, Inc.) and saved using the program of Acknowledge 3.8 (BIOPAC System, Inc.). Potentials were digitized at a sample frequency of 300 Hz and were filtered with a 60-Hz notch hardware filter and a 0.1–30 Hz band-pass filter.

During EEG recording, prime-target stimuli were presented visually in a single block of 240 trials in randomized order. Specifically, each trial consisted of (i) a prime (a Korean word) for 1000 ms, (ii) a fixation cross in the center of the computer screen in order to attract the participants’ attention to the following target stimuli, for 2000 ms, (iii) a target (English word) for a maximum of 1000 ms, during which time the participants are asked to judge the relationship between prime and target stimuli, (iv) a blank screen for 1500 ms, and finally (v) a direction: ‘NEXT WORD:’ for 500 ms. Prime words were randomly arranged from the set of 240 words (80 target words, 80 familiar words, and 80 rare words). As will be discussed later, familiar and rare words served as control. Immediately next to the offset of the target (Korean translation), participants were requested, as rapidly as possible, to press a yellow button if the prime (an English word) and target (Korean translation) matched, and to press a green button if they did not match.

4. Results

4.1. Behavioral results

In order to determine the relative efficacy of ER and PAL in long-term vocabulary retention, a 2 (exposure type: ER, PAL) × 3 (phase: pre-exposure, immediate post-exposure, delayed post-exposure) ANOVA was conducted using SPSS (since the assumption of sphericity was not violated [p = .114], no adjustments were made on the dfs for the univariate test). Results revealed a significant main effect of phase [F(2, 24) = 115.247, p = .000, partial eta squared = .906], but not of exposure type [F(1, 12) = 2.522, p = .138, partial eta squared = .174]. However, the main effect of phase was qualified by a significant interaction of phase and exposure type [F(2, 24) = 5.728, p = .009, partial eta squared = .323]. The significant interaction suggests that exposure types (ER versus PAL) do make some difference in the retention of new words, but the difference is not uniform among the different phases (pre-exposure, immediate post-exposure, delayed post-exposure).

4.2. N400 results

Next to a behavioral data analysis, we performed an EEG data analysis using MATLAB. In order to analyze ERP data, trials with electronic noise from eye and muscle artifacts were removed. The raw ERP data was referenced to an averaged reference. The data was then segmented into epochs starting from 100 ms before to 1000 ms after the onset of the stimuli. Based on visual inspection as well as previous findings [21], the ERPs of the parietal electrode
The intervention of ER and PAL showed a robust main effect of exposure type (ER versus PAL) for the parietal electrode group (Pz, P3, P4) in response to mismatched trials (t(12) = 3.176, p = .008). To determine the magnitude of the intervention (ER versus PAL), Cohen’s effect sizes showing the standardized mean difference between the mean amplitudes of two groups were calculated. The results revealed large magnitudes of exposure type effect, $d = 1.70$, for the parietal electrode group. Fig. 2 demonstrates a large N400 effect between ER and PAL groups for mismatched trials at the parietal electrode group. We interpret this ERP data to indicate that ER is more effective than PAL in light of long-term vocabulary retention.

5. Discussion and conclusion

In this study we investigated the relative effectiveness of ER and PAL in the retention of vocabulary words over short and long term. Based on converging evidence from behavioral and electrophysiological measures, we report the first evidence that both ER and PAL were effective in promoting short-term vocabulary retention and that ER was considerably more effective than PAL in long-term vocabulary retention. The findings indicate that repetition without the aid of context (i.e., PAL) may lead to short-term retention, but not to long-term retention. The finding parallels the consensus from L1 vocabulary research and some L2 research that extensive reading is a significant contributor to vocabulary acquisition [18,20]. The finding challenges the claims and findings of many L2 researchers who support decontextualized L2 vocabulary acquisition over ER [2,9,15,16,23].

As noted above, the findings of this study expand those of previous studies because it directly compared the relative efficacy of ER to PAL approaches in vocabulary building [12,20]. In addition, the novelty of data collection methods (i.e., EEG recording and extensive treatment period) and the resultant data also contribute to the current literature, because no previous studies employed such research methods.

Despite such contributions, three limitations of this study merit discussion. To begin with, knowledge of partially known words was not measured given that vocabulary learning is an incremental process: a gradual process of meeting with words, and strengthening
the small amount of knowledge gained from previous meetings [6].
In addition, the present study did not measure the working memory
capacity of participants, which is known to have a significant rela-
tionship with vocabulary learning [1]. Finally, a long prime–target
interval of 2000 ms may have eliminated some of the effects in the
immediate post-exposure phase. Future studies are warranted in
this regard.

To reiterate, the present study is the first attempt to determine the
relative efficacy of ER as opposed to PAL using both behav-
ioral and electrophysiological measures. The converging evidence
in support of ER has a great bearing on L2 pedagogy. Although
further studies are required to substantiate the initial findings of
the current study, the picture that arises from the present study is
clear: ER is more effective than PAL for long-term vocabulary reten-
tion. The finding should be considered a wake-up call for many L2
instructors and their students who tend to use PAL over ER to build
vocabulary.

Acknowledgments

The research was sponsored by KOTESOL and by the National
Research Foundation of Korea Grant funded by the Korean Govern-
ment (NRF-2010-371-A00020). Funding agencies had no additional
research roles. We would like to thank Dr. Jingu Kim and Dr. Minho
Lee for help in data acquisition and the students who generously
participated.

References

[2] A. Beaton, M. Gruneberg, N. Ellis, Retention of foreign vocabulary learned using
the keyword method: a ten-year follow-up, Second Language Research 11
(1995) 112–120.
[5] A.E. Cunningham, K.E. Stanovich, Early reading acquisition and its relation to
reading experience and ability 10 years later, Developmental Psychology 33
(1997) 934.
[6] B. Dupuy, S.D. Krashen, Incidental vocabulary acquisition in French as a foreign
[7] F.M. Hafiz, I. Tudor, Graded readers as an input medium in L2 learning, System
18 (1990) 31–42.
[9] E. Hinkel, Current perspectives on teaching the four skills, TESOL Quarterly 40
[10] M. Horst, Cobb, P. Meara, Beyond a Clockwork Orange: acquiring second lan-
guage vocabulary through reading, Reading in a Foreign Language 11 (1998)
207–233.
[12] S. Krashen, We acquire vocabulary and spelling by reading: additional evidence
[13] M. Kutas, S.A. Hillyard, Event-related brain potentials to semantically inap-
[14] B. Laufer, Second language vocabulary acquisition from language input and
learning: a case for contrastive analysis and translation, Applied Linguistics 29
[16] S. Lindstromberg, F. Boers, The mnemonic effect of noticing alliteration in lex-
[17] M.G. McKeown, The acquisition of word meaning from context by children of
high and low ability, Reading Research Quarterly 20 (1985) 482–496.
[18] W.E. Nagy, R.C. Anderson, P.A. Herman, Learning word meanings from con-
237–270.
of Research on Second Language Teaching and Learning, Lawrence Erlbaum,
in word learning reflected in event-related potentials, Journal of Experimental
[22] M. Pitts, H. White, S. Krashen, Acquiring second language vocabulary through
reading: a replication of the Clockwork Orange Study using second language acqui-
[23] D.D. Qian, ESL vocabulary acquisition: contextualization and decontextualiza-
during learning improves vocabulary acquisition: behavioral and ERP evidence,
Neuroscience Letters 482 (2010) 40–44.
[25] R. Waring, M. Takaki, At what rate do learners learn and retain new vocabu-
130–163.
[26] S. Webb, The effects receptive and productive learning of word pairs on
vocabulary knowledge, Regional English Language Centre Journal 40 (2009)
360–376.
[27] R. Zahar, T. Cobb, N. Spada, Acquiring vocabulary through reading: effects of fre-
quency and contextual richness, Canadian Modern Language Review 57 (2001)
541–572.