THIS IS AN EARLIER VERSION OF THE MANUSCRIPT. FOR THE FINAL VERSION, PLEASE CHECK THE JOURNAL WEBSITE.
Serrano, R. & Huang, H.Y. (2021). Time distribution and intentional vocabulary learning through repeated reading: a partial replication and extension. *Language Awareness*.
https://www.tandfonline.com/doi/full/10.1080/09658416.2021.1894162

Time distribution and intentional vocabulary learning through repeated reading: A partial replication and extension

Abstract

The aim of this study is to analyze the effect of different schedules of repeated reading practice on intentional vocabulary learning, and constitutes a partial replication and extension of Serrano and Huang (2018), which focused on incidental vocabulary learning. Two groups of Taiwanese EFL learners (n = 72) engaged in five repeated reading sessions; one group had the sessions on consecutive days (1-day intersession interval, ISI), whereas the other had them once a week (7-day ISI). Apart from reading for meaning, the students were also asked to focus on 36 target words. The students were tested before and immediately after the treatment. Moreover, a delayed posttest was scheduled at a retention interval (RI) of 4 and 28 days for the intensive group and spaced group respectively (considering an ISI/RI ratio of 25%). The results indicate that the short-spaced repeated reading sessions had a significantly more positive effect on vocabulary learning on both immediate and delayed posttest than the long-spaced sessions. The benefits of the short-spaced schedule were clearer for intentional vocabulary learning through repeated reading than when such learning was incidental (as in Serrano & Huang, 2018).

Key words: time distribution, lag effects, distributed-practice effects, input spacing, vocabulary learning, repeated reading

Introduction

The important role of second language (L2) practice and its distribution has been highlighted by different researchers (DeKeyser, 2007; 2017; Lightbown, 2013, Suzuki, Nakata, & DeKeyser, 2019a). In the 2010s, especially in the last few years, there has been an upsurge of interest in how time distribution of practice affects L2 learning, as reflected in many recent publications (e.g., Kasprowicz, Marsden & Sephton, 2019; Koval, 2019; Li & DeKeyser, 2019; Nakata, 2015; Nakata & Elgort, 2020; Rogers, 2017; Rogers & Cheung, 2018; 2020; Serrano & Huang, 2018; Suzuki 2017; 2018; Suzuki, 2019; Suzuki & DeKeyser, 2017a; 2017b). Despite the increase in the number of studies investigating this topic, little agreement exists as to how learning episodes should be spaced to promote higher levels of L2 learning. While some classroom-based studies provide support for distributing practice over spaced schedules (Bird, 2010; Rogers, 2015), others suggest more concentrated schedules are equally or even more effective (Kasprowicz et al., 2019; Rogers & Cheung, 2018; 2020; Serrano & Huang, 2018). Findings coming from controlled laboratory settings seem to be equally controversial, with some studies providing evidence for the positive effect of spacing repetitions of target items widely (Koval, 2019; Nakata & Webb, 2016), while others have suggested that more concentrated practice schedules might be more beneficial (Suzuki, 2017; Suzuki & DeKeyser, 2017a). Several claims have been made in the literature that some of these contradictory findings might be due to the different nature of the knowledge under investigation or learning conditions (Serrano, 2012; Suzuki & DeKeyser, 2017a).

The present study aims to examine whether distribution of practice affects intentional vocabulary learning in an equivalent way as incidental learning (Serrano & Huang, 2018). In cognitive psychology, differential spacing effects have been found for these two learning

conditions (Toppino & Bloom, 2002; Verkoeijen, Rikers, & Schmidt, 2005), and thus it is pertinent to analyze these effects for L2 learning.

Replication research is crucial in order to advance scientific research and should be encouraged in SLA to further extend and test the generalizability of previous findings (Marsden, Morgan-Short, Thompson, & Abugaber, 2018; Porte & McManus, 2019). Considering the lack of consistent results for how spacing affects L2 learning, replication studies are an obvious need. By performing conceptual replications of their own studies, Rogers and Cheung (2020) and Suzuki (2017) could further consider their initial results and throw more light on distributed-practice effects. The present study, which is a partial replication and extension of Serrano and Huang (2018), also hopes to further examine previously reported findings on vocabulary learning through repeated reading under different schedules and, most importantly, contribute to the debate concerning potentially differential distribution of practice effects for different L2 learning conditions. The design, methodology and analyses of the original study were closely followed, with the exception of the instructional approach, which, in the initial study encouraged incidental learning, but in the current study promoted intentional learning.

Literature review

Distributed-practice effects in cognitive psychology

The term *distributed-practice effects* has been used as a generic term that includes both the spacing and lag effects (Toppino & Gerbier, 2014). The *spacing effect* is one of the most robust findings in the cognitive psychology literature (Cepeda, Pashler, Vul, Wixted, & Rohrer, 2006). It suggests that items that appear in spaced sequences (with time or other intervening items in between) are better learned and remembered than items in massed

sequences (immediate repetitions). Similarly, when considering distributed practice, research suggests that longer spacing seems to lead to more learning than shorter spacing (*lag effect*). Although the spacing effect has been found consistently, lag effects are not so uniform (Rogers, 2017; Toppino & Gerbier, 2014).

There are different theories that have tried to account for distributed-practice effects. Deficient processing theories suggest that massed repetitions are not fully processed because previous memory traces are too accessible (Glanzer, 1969; Hintzman, 1976). Encoding variability hypotheses suggest that spaced repetitions benefit from different contextual cues, as they appear in different contexts (vs. massed), and such contextual cues help future retrieval (Glenberg, 1979).

One interesting finding concerning distributed-practice effects is that more spacing is more conducive to learning up to a point in which learning starts to decrease when spacing increases (Cepeda et al., 2006). Study phase retrieval theories (Toppino & Bloom, 2002) claim that spacing makes retrieval more effortful, which, following Bjork's desirable difficulty framework (Bjork, 1994), could be positive for learning (Suzuki, Nakata & DeKeyser, 2019b). Nevertheless, if spacing is too wide, it can lead to failure to retrieve previous presentations of the target items. Bahrick and Phelps (1987) claim that "The optimum interval is likely to be the longest interval that avoids retrieval failures" (p. 349).

Another factor to consider in the investigation of distributed-practice effects is the retention interval between the end of the learning phase and the testing phase. Many studies have found that, at short retention intervals, distributed-practice effects do not occur (Bahrick & Hall, 2005; Pavlik & Anderson, 2005). Cepeda et al. (2006) claim that retention increases as intervals between learning sessions increase. Rohrer and Pashler (2007) further suggest that the ideal intersession interval (ISI) should be between 10-30% of the retention interval

(RI). For instance, considering two learning conditions, one in which there is a 1-day ISI (the treatment happens on consecutive days) and another in which there is a 7-day ISI (i.e., once a week), learners in the former group are expected to outperform those in the latter at a test scheduled 10 days after the treatment (the ratio ISI/RI = 10% for 1-day ISI, but 70% for the 7-day ISI). The opposite results would be expected if the test was performed at a 30-day RI (ratio ISI/RI = 23.3% for 7-day ISI but 3.33% for 1-day ISI).

Although research shows that distributed-practice effects happen both in the case of intentional and incidental learning (Toppino & Gerbier, 2014), Verkoeijen et al. (2005) claim that incidental learning conditions start being negatively affected by longer lags sooner than intentional conditions. In their study, the spacing effect was found for word learning when considering inter-repetition lags of 0, 2 and 5 intervening words in both incidental and intentional learning. However, the spacing effect disappeared after 8 intervening words in the case of incidental learning, but only after 14 words for intentional learning, suggesting that intentional learning is less subject to memory decay due to spacing than incidental learning.

Distributed-practice effects in L2 vocabulary learning

In the studies that have analyzed learning of discrete, decontextualized L2 lexical items in experimental conditions, there is robust evidence that spaced items are better learned and remembered than massed (Nakata, 2015). When considering spaced conditions, however, there is not clear evidence that more spacing leads to higher learning than less spacing (Nakata & Webb, 2016). In a similar vein, recent studies in classroom settings have failed to find evidence in support of distributed-practice effects. In a primary school in Hong-Kong, Rogers and Cheung (2018) analyzed learning of words that were repeated in spaced-

short (1 day ISI) vs. spaced-long (8-day ISI) lags. A test performed four weeks after the treatment suggests that spaced-short items were better learned than spaced-long items, contrary to lag effects. Their replication study did not find significant differences between the two conditions (Rogers & Cheung, 2020). Küpper-Tetzel, Erdfelder and Dickhaüser (2014) examined English vocabulary learning in the case of German high school students. The repetitions of the target words occurred in one session (massed), or two sessions spaced over a 1-day ISI or over a 10-day ISI. The results suggest an advantage for the 1-day ISI at a 7-day RI, while both 1-day and 10-day ISI obtained better results than massed at a 35-day RI. These results further support the importance of considering the retention interval when examining distributed-practice effects.

Distributed-practice effects in vocabulary learning through reading

Reading has been shown to be an excellent source of vocabulary learning (Horst, 2005; Pellicer-Sánchez, 2016; Webb & Chang, 2012). Some of the studies focusing on incidental vocabulary learning through reading also examined the effect of the distribution of repeated exposures to the same words in the texts, as a possible factor affecting vocabulary learning (even though they were not originally designed to analyze this variable). Elgort and Warren (2014) found that, when the spacing between repetitions was too wide (not appearing within the same chapter), learners with lower proficiency failed to learn those words, possibly because of the decay of previous episodic memory traces. On the other hand, in the study by Elgort, Brysbaert, Stevens and Van Assche (2018), words that appeared over two reading sessions were better learned than those appearing only during one reading session. In Webb and Chang (2015), the distribution of repeated occurrences of the target words over 10 graded readers did not predict vocabulary gains.

The studies by Koval (2019) and Nakata and Elgort (2020) focused on distributed-practice effects on vocabulary learning through sentence reading in experimental conditions. In both studies the target words (Finnish words and pseudo-words respectively) appeared in English sentences that were either massed (immediate repetitions) or spaced. Both studies found support for the spacing effect in explicit word learning.

To the author's knowledge, Serrano and Huang (2018) is the only study that has examined distributed-practice effects on vocabulary learning through repeated reading of a text in a classroom setting. Repeated reading involves reading short texts multiple times, which has been shown to promote reading fluency and comprehension (Chang & Millet, 2013) and also encourage L2 incidental vocabulary learning (Han & Chen, 2010; Liu & Todd, 2016; Webb & Chang, 2012). One critical aspect to consider regarding the implementation of repeated reading is how repetitions should be spaced to maximize vocabulary learning. Serrano and Huang (2018) examined two schedules of five repeated reading sessions (with audio support), which they referred to as intensive (1-day ISI) and spaced (7-day ISI). Their participants were Taiwanese learners of English in grade 10. They read the text for general comprehension (after each reading, the students answered several comprehension questions, but there was no focus on the target vocabulary). The text included 36 target words, most of which were expected to be unknown. The results of an L2-L1 vocabulary matching test suggest that the intensive sessions were more conducive to incidental vocabulary learning at a short-term RI (immediate posttest, right after the treatment). Moreover, although there were no significant differences for long-term retention in a delayed posttest between the two groups, the intensive group experienced significant loss of the acquired vocabulary from post- to delayed posttest, whereas this was not the case for the spaced group.

Serrano and Huang (2018) considered their treatment incidental because the participants were asked to read for general comprehension and they were not forewarned of a subsequent vocabulary test (Hulstijn, 2003; 2005). This operationalization of incidental learning has been widely adopted in vocabulary studies (Peters, Hulstijn, Sercu & Lutjeharms, 2009), as it is easier for researchers to control testing conditions than participants' orientation towards learning. The current study will follow the same operationalization based on experimental conditions, according to which the present treatment will be considered intentional because the students were explicitly instructed to learn the target vocabulary and they also knew that there would be a vocabulary test after the reading sessions. The difference between the treatment in Serrano and Huang (2018) and the current study is thus the focus that vocabulary received in the teaching approach and in test announcement: no explicit instructions to focus on vocabulary learning (original study) vs. explicit instructions to learn the target words for a subsequent vocabulary test (current study).

Although research has shown that L2 learners can incidentally increase their vocabulary through reading, it has also been found that when learners are required to pay attention to the target words (e.g., by vocabulary test announcement, by doing post-reading vocabulary activities, etc.), they make more lexical gains than under purely incidental "reading-only" conditions (Paribakht & Wesche, 1997, Peters et al., 2009; Sonbul & Schmitt, 2010). The purpose of the current study is to examine in which way creating intentional learning conditions has an effect on vocabulary learning through repeated reading under different schedules.

The present study

The present study aims to fill important gaps in the SLA literature. First, it will examine to what extent repeated reading can promote vocabulary gains in intentional learning conditions (vs. the typically investigated incidental conditions). Additionally, it will explore whether the distribution of the sessions has an effect on intentional vocabulary learning through repeated reading. Crucially, by conducting a partial replication and extension of Serrano and Huang (2018), the present study will examine whether intentional and incidental learning conditions are affected by spacing in similar or different ways. The study will answer the following questions, which are equivalent to the ones in the initial study:

1. To what extent can EFL learners learn new vocabulary from repeated reading when they are explicitly instructed to focus on vocabulary learning as well as on reading comprehension?
2. Does time distribution of repeated reading sessions (intensive vs. spaced) affect the way vocabulary is learned and retained under intentional learning conditions?

Considering previous research, we expect significant vocabulary gains as a consequence of repeated reading with a focus on vocabulary learning (Han & Chen, 2010). In line with previous vocabulary studies in classroom settings, we hypothesize that the intensive schedule will promote as much as, or more learning and retention than the spaced schedule (Rogers & Cheung, 2018; 2020; Serrano & Huang, 2018).

Methods

Participants

The participants of this study included 72 Taiwanese learners of English in two intact grade 10 classes of 36 students each, who were 15-16 years old. All the students had been

learning English for three years, and were currently receiving 5 h of English instruction per week. One of these two groups was randomly assigned to the intensive repeated reading condition and the other one to the spaced condition.

Previous vocabulary knowledge influences reading fluency and comprehension (Jeon & Yamashita, 2014; Maun, 2009). In order to make sure that the two groups had similar previous vocabulary knowledge, all the students performed the same test as in the initial study: a bilingual English-Chinese Vocabulary Levels Test (VLT) (Webb & Chang, 2012) up to the 3K word level, including 90 multiple-choice questions. As in previous studies using the test, the scores obtained in each word frequency band were added and then multiplied by 33.3 to get a final score, which provides an approximate reflection of students' vocabulary size (Chang & Millet, 2013; Webb & Chang, 2015). When analyzing the scores of the test, it became apparent that there were three outliers in the intensive group, who obtained a score below 1,000 (231, 627 and 792), while the average was score 1,977 (*SD* = 341.77) for the intensive group and 1,927.58 (*SD* = 295.39) for the spaced group. These average scores were quite similar to the scores obtained by the participants in Serrano and Huang (2018), ensuring the comparability of the two pools of participants: 1,921.14 (*SD* = 356.9) and 1,957.68 (*SD* = 265.6) respectively. Because of their low previous vocabulary knowledge, the three outliers were eliminated from the study, which finally consisted of 33 students in the intensive group (26 females) and 36 in the spaced group (26 females). There were no significant differences between the two groups in the scores of the English-Chinese VLT once these outliers were removed: $t(63) = .636$, $p = .527$.

The two classes were taught by the same teacher, who followed the same procedure in administering the treatment, which encouraged intentional vocabulary learning through repeated reading of the same text.

Reading passage and target words

The reading passage, with the title *Kyoto: The Spiritual Heart*, was the same as in the original study. It was a short text (419 words) about the Japanese city and its culture, and it included 36 target words. As was the case in Serrano and Huang (2018), the participants in the current study were expected to know some of the target vocabulary (especially the words in the 1K and some in the 2K frequency band). Including known words in vocabulary tests is expected to have an encouraging effect for the students (Webb & Chang, 2015). See Appendix A for the list of words and their frequency from the BNC/COCA word frequency lists, as retrieved from Compleat Web VP, LexTutor (Cobb, n.d.).

Treatment

The treatment consisted of five repeated reading sessions in which the students read and listened to an aural rendition of the same text about Kyoto (4 min.). As in Serrano and Huang (2018), the treatment concerns repeated reading-while-listening, or assisted repeated reading. However, in order to make the reading of the present manuscript less cumbersome, the terminology used for the treatment is simply repeated reading. The students received the reading passage and the teacher asked them to read and listen to it at the same time in order to understand the main ideas, while also paying special attention to the highlighted vocabulary (target words). After reading-while-listening to the text once, in each session the students answered a different set of 5-6 comprehension questions (including fill-in-the blank, multiple-choice, and true-false questions), which asked about general or specific ideas developed in the reading passage. All the comprehension questions encouraged students to process the target vocabulary; however, there were no explicit vocabulary-focused activities.

The students received the comprehension questions (which were the same as in the initial study) together with a glossary on a separate sheet where the target words appeared together with their Chinese translations. Apart from answering the comprehension questions, the students were instructed to learn the meaning of the target words for a subsequent test (the exact date of which was not announced). The sessions lasted for 20 minutes and after each session, the teachers collected all the material (reading passage, glossary and comprehension questions).

Vocabulary test

The participants' knowledge of the target vocabulary was assessed through a bilingual matching vocabulary test based on the one used by Webb and Chang (2015) (same as in Serrano & Huang, 2018). The 36 target words were distributed in three blocks of 12. Two of the blocks included only nouns and one included six verbs and six adjectives. Each block contained the target words in English and the corresponding translations into Mandarin Chinese plus two distractors. The target items received a score of 1 when they were correctly matched with their Chinese translation and 0 when it was not the case. The test had good reliability (Cronbach's alpha = .856).

Procedure

The study included a pretest, treatment, immediate and delayed posttest, and was performed over seven different sessions. Two days before the treatment started, all the students did the pretest including the 36 target words (15 minutes). On the first day of the treatment, the students first performed the VLT (which lasted for about 20 minutes), and then the first reading session (20 minutes). The four subsequent readings were differently spaced

for intensive vs. spaced groups: the former group did the sessions on four consecutive days (1 day ISI), whereas the latter did it once a week over four weeks (7-day ISI). On the last day of the treatment, after the students performed the last reading of the passage and the comprehension questions, they did the target vocabulary immediate posttest (see Figure 1 for a visual representation of the procedure). As in Serrano and Huang (2018), and following Rohrer and Pashler (2007), the delayed posttest was scheduled considering an ISI/RI ratio of 25%, and consequently happened 4 and 28 days after the end of the treatment for the intensive group and spaced group respectively. Although this is a limitation of the original study, the same RIs were kept because the aim of the current study is to isolate the influence of intentional vs. incidental learning while keeping all the other variables constant.

[Figure 1]

Statistical analyses

In order to facilitate comparisons with Serrano and Huang (2018), the same statistical analyses were performed. First, students' learning of the target vocabulary in the two conditions (intensive vs. spaced) was examined through a Repeated Measures ANOVA, after confirming that the data were normally distributed. Condition (intensive vs. spaced) was the between-subject factor, whereas time (pretest, immediate posttest, delayed posttest) was the within-subject factor. The interaction between the two was also examined. When significant results were obtained, pairwise comparisons were performed after applying the Bonferroni adjustment for multiple comparisons. Additionally, independent-sample *t*-tests were also performed to compare *gains* by condition, and paired-sample *t*-tests to analyze gains within each condition separately.

When interpreting effect sizes, the recommendations by Plonsky and Oswald (2014) are followed: d = .40 (small); d = .70 (medium) and d = 1.0 (large) for between-group comparisons and d = .60 (small), d = 1.0 (medium) and d = 1.40 (large) for within-group comparisons. For partial eta squared: small = .0099; medium = .0588; large = .1379 (Cohen, 1988).

Results

The overall results of vocabulary learning through reading in this study are quite positive (see Table 1). There were higher scores in the immediate posttest than in the pretest for both conditions, although the gains were higher in the intensive group. Additionally, the scores of the delayed posttest were slightly lower than in the immediate posttest for both conditions.

[Table 1]

The results of the Repeated Measures ANOVA, with condition as the between-subject factor and testing time as the within-subject factor suggest that there was a significant effect of time with a large effect size: $F(2, 134) = 198.62$, $p < .001$, partial eta squared = .748. Pair-wise comparisons show significant differences ($p < .001$) between all testing times, indicating overall significant short- and long-term learning and also significant forgetting. The effect sizes were large for pre- to immediate posttest (d = -1.64), medium (d = 1.15) from pre- to delayed posttest, and small from immediate to delayed posttest (d = 0.35). The effect of condition was also significant, in favor of the intensive group (medium effect size): $F(1, 67)$

= 5.56, p = .021, partial eta squared = .077. Similarly, there was a significant interaction between time and condition, also with a medium effect size: $F(2, 134) = 9.37$, $p < .001$, partial eta squared = .123 (see Figure 2).

[Figure 2]

The results so far parallel those reported by Serrano and Huang (2018), except for the effect of condition, which was not significant in the original study. The effect size was equivalent for time, but it was large for the interaction (while it was medium in the current study). In order to explore further this interaction other between- and within-group analyses were performed.

Three *t*-tests with the gain (immediate posttest minus pretest; delayed posttest minus pretest) and retention scores (delayed minus immediate posttest) were performed. After the Bonferroni adjustment, the alpha level was set to .0166. These *t*-tests show that the immediate gains (from pre- to immediate posttest) were significantly higher for the intensive condition (M = 11.48, SD = 4.79) than for the spaced (M = 8.61, SD = 4.85): $t(67) = 2.47$, p = .016, and the effect size was small (d = 0.59). Retention scores were higher, but not significantly, for intensive (M = -1.75, SD = 2.83) than for spaced (M = -3.36, SD = 4.01) repeated reading: $t(62.97) = 1.93$, p = .058, and the effect size was small (d = 0.46). Finally, gains from pre- to delayed posttest were also significantly higher in the intensive (M = 9.72, SD = 4.67) than in the spaced condition (M = 5.25, SD = 4.54): $t(67) = 4.03$, $p < .001$, with a close to large effect size: d = 0.97. These results only partly confirm those of the initial study, which found that 1) immediate gains were significantly higher for the intensive group (small effect size); 2) retention was significantly higher for the spaced group (large effect size); 3) there were no

significant differences between conditions for long-term learning. Table 2 provides a summary of the findings of the two studies considering gains by condition.

[Table 2]

Within-group comparisons through *t*-tests (alpha level set to .008 after applying the Bonferroni adjustment) indicated similar behavior in the intensive and the spaced groups, as both significantly learned from the treatment with large effect sizes (intensive $t(32) = -13.77$, $p < .001$, $d = -1.95$; spaced $t(35) = -10.65$, $p < .001$, $d = -1.42$), both experienced significant losses, although the effect size was smaller for the intensive group (intensive: $t(32) = 3.57$, $p = .001$, $d = 0.24$; spaced: $t(35) = 5.02$, $p < .001$, $d = 0.50$), and both groups also made long-term gains, with a large effect size in the intensive group and small in the spaced (intensive: $t(32) = -11.94$, $p < .001$ $d = -1.65$; spaced: $t(35) = -6.93$, $p < .001$, $d = -0.86$). Although the results are quite similar for the two groups in terms of *p* values, the effect sizes suggest that there were comparable immediate gains, but more loss for the spaced from immediate post- to delayed posttest. Also, the effect size for long-term gains is large for the intensive and small for the spaced group. These results only partly replicate the results reported in the incidental learning study: 1) both groups made significant gains from pre- to immediate posttest (medium effect size for intensive, small for spaced); 2) there was a significant loss for the intensive group (medium effect size), but not for the spaced; 3) there were significant gains from pre- to delayed posttest in both conditions (although the effect size was larger in the spaced group). The most salient difference between the two studies is that the losses for the spaced group were not significant in the incidental learning study, but they were in the present study (see Table 3 for a summary).

[Table 3]

If we examine gains in terms of percentage of words learned out of the words that could still be learned (considering that pretest scores on average were around 13/36), the students in the intensive group learned 52.11% of the words they could still learn (11.48 out of 22.03) from pre- to immediate posttest, lost 15.33% and learned 44.12% from pre- to delayed posttest. For the spaced condition, there was a gain of 37.51% from pre- to immediate posttest (8.61 out of 22.95), a loss of 15.51%, and an overall gain of 22.87% from pre- to delayed. Table 4 shows the gains in terms of raw number of words and percentage of words learned in parentheses. For ease of comparison, the results of Serrano and Huang (2018) are also presented.

As can be seen in Table 4, the learners in the current study learned more words both at a short retention interval (pre- immediate posttest) and at a long retention interval (pre-delayed posttest) than in Serrano and Huang (2018). However, it seems like intentional learning was more effective for the intensive than for the spaced group in the long term. While for the spaced groups there are practically no differences between intentional and incidental vocabulary learning through repeated reading at a long retention interval (gains pre-delayed), it seems like the intensive condition benefits more from an intentional approach for long-term learning.

[Table 4]

Discussion

The present study set out to explore how time distribution of repeated reading affects intentional vocabulary learning. The study is a partial replication and extension of Serrano and Huang's (2018) study on incidental vocabulary learning. This type of replication is crucial in order to observe differential effects of spacing for different types of learning conditions or instructional approaches, while keeping other variables constant. The context, design, methodology and analyses were the same as in the original study; the only factor that changed was the degree of focus on the learning of the target vocabulary. Whereas in Serrano and Huang (2018) the focus was exclusively directed towards reading comprehension, the participants in the current study were also asked to learn the target vocabulary in order to perform a vocabulary test. In both studies, vocabulary learning took place in meaningful contexts while the students were reading or answering reading comprehension questions and there were no specific vocabulary exercises in any of the two studies. This type of learning is different from the type of focused learning approaches used in experimental studies of the spacing effect for vocabulary acquisition (through pair-associate learning, for example), where rote learning of isolated words (or word pairs) is examined (e.g., Nakata, 2015). It is also different from previous experimental studies on distributed-practice effects on vocabulary learning through reading (e.g., Koval, 2019; Nakata & Elgort, 2020), in that it involves reading of a whole text repeatedly (vs. sentence reading in which words appear in different contexts). Crucially, the present study is a classroom-based study, which differs from previous classroom studies on vocabulary learning (Küpper-Tetzel et al., 2014; Rogers & Cheung, 2018) in that it involves learning of words in context and not in isolation.

The first research question asked to what extent vocabulary could be learned intentionally from reading the same text repeatedly. The results of the analyses indicated that,

regardless of spacing condition, there were significant gains immediately after the treatment and at a longer retention interval in which ISI = 25% of RI. The effect size of these differences was large in the pre-immediate posttest comparison and medium in the pre-delayed posttest comparison. These results are in line with Serrano and Huang (2018), except that the effect size was medium immediately after the treatment and small at a longer RI. These results might suggest that making students focus on target vocabulary is slightly more effective for subsequent vocabulary learning than a pure focus on reading comprehension, in line with other studies (Paribakht & Wesche, 1997). Intentional learning makes stronger memory representations across repetitions (Verkoeijen et al., 2005), which leads to better performance on vocabulary tests.

The second research question focused on differences between intensive and spaced repeated reading for intentional vocabulary learning. The results suggest that 1) the intensive condition was significantly more beneficial for short-term learning; 2) the intensive group experienced less forgetting; and 3) the intensive group made significantly more gains from pre- to delayed posttest.

The benefits of the most concentrated schedule for short-term learning are in line with claims by Cepeda et al. (2006) which suggest that more concentrated study sessions make the target material more easily retrievable when testing happens within short RIs. Similarly, these results are in agreement with other vocabulary studies on distributed-practice effects (although the RIs vary). Küpper-Tetzel et al. (2014) considered two RIs: a shorter one of 7 days and a longer one including 35 days. The authors found an advantage for the 1-day ISI condition vs. 10-day ISI for their short RI.

In the current study, the results concerning the advantage of the intensive treatment for long-term learning are in line with those reported by Rogers and Cheung (2018), despite

the difference in RIs between the two studies. The advantage for the intensive group might come from the shorter lag between immediate and delayed posttest. Nevertheless, it should be emphasized that, in similar conditions (i.e., with a shorter RI), the intensive group did *not* outperform the spaced group in Serrano and Huang (2018). One explanation for the results of the current study might be that the lag between reading sessions in the spaced condition (7-day ISI) might be too long for previous memory traces of the target words to remain active in participants' memory. Consequently, retrieval of previous presentations when repetitions appear in the spaced condition might be unsuccessful, limiting the effect of extra-exposures. Comparing the results of the spaced group in the current study and in Serrano and Huang (2018), it is unexpected that there were more losses in the intentional than in the incidental condition from immediate to delayed posttest (see Table 3). However, it should be noted that the amount of learning in the spaced incidental group was the lowest of all the conditions (which might partly explain the higher retention rate), and that the scores of the delayed posttest were quite similar in the two studies for this group (see Table 4).

The findings of the current study are only partly in line with the original study of incidental learning. The only common result refers to immediate vocabulary gains, which, in both studies, are higher for the intensive condition. Nevertheless, the other results are different, as in Serrano and Huang (2018) the intensive group experienced more forgetting from immediate to delayed posttest and there were no significant differences between conditions at a longer RI (even though descriptively the spaced group experienced more learning). The findings concerning forgetting in Serrano and Huang (2018) show that for incidental learning the significant gains in the intensive condition in the immediate posttest might have been a reflection of recency effects more than actual learning: in the delayed posttest, the intensive group did not show any evidence of knowledge of more than half of

the words that had been "learned". In the intentional condition in the present study, however, the automatic decrease in attention expected for the intensive schedule (which can prevent memory consolidation, in line with deficient processing theories) might have been counteracted by a voluntary effort in word learning coming from the treatment (the instructions asked the students to answer comprehension questions as well as to learn the target vocabulary). This voluntary effort in learning might have enhanced not only short-term but also long-term memory representations of the target words in the intentional intensive condition. Overall, and considering Serrano and Huang (2018), the results of the present study suggest that concentrated exposure to target words through repeated reading enhances word learning more significantly when there is an additional focus on word learning apart from general comprehension.

Suggestions for future (replication) research

Being a partial replication, the present study acknowledges the same limitations as were mentioned by the authors of the original study. One of the main limitations is that, although the ratio ISI-RI is the same for intensive and spaced conditions (25%), the lag between the end of the treatment and the delayed posttest is necessarily longer for the spaced (28 days) than for the intensive group (4 days). Because of this, while it is safe to make between-group comparisons in the immediate posttest, the comparisons in the delayed posttest should be interpreted with caution. Despite this limitation, the interesting finding from the current study in relation to Serrano and Huang (2018) is that the same RIs led to different results for intentional and incidental learning. Intentional learning conditions with explicit vocabulary focus encouraged more learning and retention when repeated reading happened intensively. On the contrary, incidental learning conditions appeared to be more

beneficial for the spaced group in terms of retention (despite the longer lag between immediate post- and delayed posttest).

Future studies, however, should address the limitation mentioned above. One option would be to test the two groups at the two RIs, which would involve testing the students four times (pre-, immediate, short-RI, long-RI). One obvious disadvantage of such design is that more opportunities for task-repetition effects would be created. Other alternatives to minimize those effects would be testing half of the participants in each group at a short-RI and the other half at a long-RI, or avoid immediate testing and only use long-term learning at the same RI for both groups.

Another limitation is that it is not possible to isolate the unique contribution of repeated reading to the learning of the target words, as the treatments included comprehension questions that invited learners to revise the text and probably process it, and the target words, more deeply than if they were just reading. Future studies could eliminate these comprehension questions in order to identify more clearly the specific role of repeated reading, although such as study would be more difficult to integrate in a classroom setting.

Additionally, as in the original study, the target words in the selected reading passage included some words borrowed from Japanese (geisha, kimono, omamori, Zen), which might have been learned and processed differently from the other English words. Future studies should also consider this.

Finally, another limitation of the current study is that the learners might have engaged in intentional, even discrete-item learning outside of the repeated reading sessions, due to the nature of the treatment itself (vs. the incidental treatment), although such learning was not facilitated because all the materials were collected after each session.

Conclusion

In conclusion, the results of the study in combination to those of the initial study by Serrano and Huang (2018) show that the effect of spacing might be different for different learning conditions or instructional approaches and go in line with previous claims that distributed-practice effects might vary depending on a number of factors, such as the type of knowledge under investigation (Li & DeKeyser, 2019; Nakata & Elgort, 2020). More replication studies of existing research should be performed, in order to examine the effect of time distribution of L2 learning practice under both the same and different conditions in order to generalize and extend previous findings. As suggested in the literature review, the findings regarding lag effects are especially controversial in the SLA literature (vs. more uniform findings for the spacing effect that focus on massed vs. spaced sequences).

One possible pedagogical implication that can be derived from our findings, together with those from the original study, is that, if short-term vocabulary learning is to be maximized, it would be better for teachers to organize repeated reading sessions within short-intervals. In order to encourage retention of vocabulary learned under an intensive schedule, it would be better to create intentional learning conditions by asking learners to focus on the target vocabulary. On the other hand, teachers can also include repeated reading and/or reading-while-listening in their classes to practice other skills, such as reading fluency, listening and reading comprehension, and expect learners to acquire vocabulary incidentally from the activity, and even remember it after a few weeks if done under a spaced schedule. In all different conditions, however, repeated reading in combination with other activities seems to promote vocabulary learning, and it can certainly be considered as an additional way for learners to develop their lexicon. More research should be performed in order to generalize these findings to other participant groups and other ISIs and RIs.

References

Bahrick, H. P. and Hall, L. K. (2005). The importance of retrieval failures to long-term retention: A metacognitive explanation of the spacing effect. *Journal of Memory and Language, 52,* 566–577.

Bahrick H. P. and Phelps, E. (1987) Retention of Spanish vocabulary over 8 years. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 13, 344-349.

Bird, S. (2010). Effects of distributed practice on the acquisition of second language English syntax. *Applied Psycholinguistics, 31,* 635–650.

Bjork, R. A. (1994). Memory and metamemory considerations in the training of human beings. In J. Metcalfe & A. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 185–205). MIT Press.

Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin, 132,* 354–380.

Chang, C. S., & Millett, S. (2013). Improving reading rates and comprehension through timed repeated reading. *Reading in a Foreign Language Journal, 25,* 126-148.

Cobb, T (n.d.). Lextutor. Vocabprofile [computer program]. Accessed at https://www.lextutor.ca/vp/comp/.

Cohen, J. (1988). Statistical power analysis for the behavioural sciences (2nd ed.). Academic Press: New York.

DeKeyser, R.M. (2007) Introduction: Situating the concept of practice. In R.M. DeKeyser (Ed.), *Practice in second language: Perspectives from applied linguistics and cognitive psychology* (pp. 1-18). Cambridge: CUP.

DeKeyser, R. (2017). Knowledge and skill in ISLA. In S. Loewen and M. Sato (Eds.), *Routledge handbook of instructed second language acquisition* (pp. 15-32). London: Routledge.

Elgort I. & Warren P. (2014). L2 vocabulary learning from reading: Explicit and tacit lexical knowledge and the role of learner and item variables. *Language Learning, 64*, 365–414.

Elgort, I., Brysbaert, M., & Stevens, M., & Van Assche, E. (2018). Contextual word learning during reading in a second language: An eye-movement study. *Studies in Second Language Acquisition, 40*, 341–366.

Glanzer, M. (1969). Distance between related words in free recall: Trace of the STS. *Journal of Verbal Learning and Verbal Behavior, 8*, 105–111.

Glenberg, A. M. (1979) Component-levels theory of the effects of spacing of repetitions on recall and recognition. *Memory & Cognition, 7*, 95-112.

Han, Z., & Chen, C. A. (2010). Repeated-reading-based instructional strategy and vocabulary acquisition : A case study of a heritage speaker of Chinese. *Reading in a Foreign Language, 22*, 242–262.

Hintzman, D. L. (1976) Repetition and memory. In G. H. Bower (Ed.), *The psychology of learning and memory* (pp. 47-91). New York: Academic Press.

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

Hulstijn, J. H. (2003). Incidental and intentional learning. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 349-381). Oxford: Blackwell.

Hulstijn, J. H. (2005). Theoretical and empirical issues in the study of implicit and explicit second language learning: Introduction. *Studies in Second Language Acquisition, 27*, 129‑140.

Jeon, E. H., & Yamashita, J. (2014). L2 reading comprehension and its correlates: A meta-analysis. *Language Learning, 64*, 160-212.

Kasprowicz, R. E., Mardsen, E., & Sephton, N. (2019). Investigating distribution of practice effects for the learning of foreign language verb morphology in the young learner classroom. *The Modern Language Journal, 103,* 580-606.

Koval, N. G. (2019). Testing the deficient processing account of the spacing effect in second language vocabulary learning: Evidence from eye tracking. *Applied Psycholinguistics, 40*, 1103-1139.

Küpper-Tetzel, C.E., Erdfelder, E., & Dickhäuser, O. (2014). The lag effect in secondary school classrooms: Enhancing students' memory for vocabulary. *Instructional Science, 42*, 373‑388.

Li, M., & DeKeyser, R. (2019). Distribution of practice effects in the acquisition and retention of L2 Mandarin tonal word production. *The Modern Language Journal, 103,* 607-628.

Lightbown, P. (2013). Making the minutes count in L2 teaching. *Language Awareness, 23*, 3-23 .

Liu, Y. T., & Todd, A. G. (2016). Implementation of assisted repeated reading techniques for the incidental acquisition of novel foreign vocabulary. *Language Teaching Research, 20*, 53-74.

Marsden, E., Morgan-Short, K., Thompson, S. & Abugaber, D. (2018). Replication in second language research: Narrative and systematic reviews and recommendations for the field. *Language Learning, 68,* 321-391.

Maun, I. (2009). Scaffolds for reading in French: Lessons from history, guidance for the future? *Language Awareness, 18,* 198-214.

Nakata, T. (2015). Effects of expanding and equal spacing on second language vocabulary learning. *Studies in Second Language Acquisition, 37,* 677–711.

Nakata, T. & Elgort, I. (2020). Effects of spacing on contextual vocabulary learning: Spacing facilitates the acquisition of explicit, but not tacit, vocabulary knowledge. *Second Language Research.* Online first.

Nakata, T., & Webb, S. (2016). Does studying vocabulary in smaller sets increase learning? The effects of part and whole learning on second language vocabulary acquisition. *Studies in Second Language Acquisition, 38,* 523-552.

Paribakht, T. S., & Wesche, M. (1997). Vocabulary enhancement activities and reading for meaning in second language vocabulary development. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition: A rationale for pedagogy* (pp. 174–200). New York: Cambridge University Press.

Pavlik, P. I. and Anderson, J. R. (2005). Practice and forgetting effects on vocabulary memory: An activation-based model of the spacing effect. *Cognitive Science, 29,* 559-586.

Pellicer-Sánchez, A. (2016). Incidental vocabulary acquisition from and while reading: an eye-tracking study. *Studies in Second Language Acquisition, 38,* 97-130.

Peters, E., Hulstijn, J. H., Sercu, L., Lutjeharms, M. (2009). Learning L2 German vocabulary through reading: The effect of three enhancement techniques compared. *Language Learning, 59,* 113-151.

Plonsky, L., & Oswald, F. L. (2014). How big is 'big'? Interpreting effect sizes in L2 research. *Language Learning, 64,* 878-912. http://dx.doi.org/10.1111/lang.12079

Porte, G., & McManus, K. (2019). *Doing replication research in applied linguistics*. New York and London: Routledge.

Rogers, J. (2015). Learning second language syntax under massed and distributed conditions. *TESOL Quarterly, 49,* 857-866.

Rogers, J. (2017). The spacing effect and its relevance to second language acquisition. *Applied Linguistics, 38,* 906-911.

Rogers, J., & Cheung, A. (2018). Input spacing and the learning of L2 vocabulary in a classroom context. *Language Teaching Research*. Advance online publication.

Rogers, J. & Cheung, A. (2020). Does it matter when you review? Input spacing, ecological validity, and the learning of L2 vocabulary. *Studies in Second Language Acquisition* FirstView articles.

Rohrer, D., & Pashler, H. (2007). Increasing retention without increasing study time. *Psychological Science, 16*(4), 183–186.

Serrano, R., & Huang, H. (2018). Learning vocabulary through assisted repeated reading: How much time should there be between repetitions of the same text? *TESOL Quarterly, 52,* 971-994.

Sonbul, S., & Schmitt, N. (2010). Direct teaching of vocabulary after reading: Is it worth the effort? *ELT Journal, 64,* 253-260.

Suzuki, Y. (2017). The optimal distribution of practice for the acquisition of L2 morphology: A conceptual replication and extension. *Language Learning, 67,* 512-545.

Suzuki, Y. (2018). The role of procedural learning ability in automatization of L2 morphology under different learning schedules: An exploratory study. *Studies in Second Language Acquisition. 40,* 923-937.

Suzuki, Y., & DeKeyser, R. (2017a). Effects of distributed practice on the proceduralization of morphology. *Language Teaching Research,21,* 166-188.

Suzuki, Y., & DeKeyser, R. (2017b). Exploratory research on L2 practice distribution: An aptitude-by-treatment interaction. *Applied Psycholinguistics, 38,* 27-56.

Suzuki, Y., Nakata, T., DeKeyser, R. (2019a). Optimizing second language practice in the classroom: Perspectives from cognitive psychology. *The Modern Language Journal, 103,* 551-561.

Suzuki, Y., Nakata, T., & DeKeyser, R. M. (2019b). The desirable difficulty framework as a theoretical foundation for optimizing and researching second language practice. *Modern Language Journal, 103*(3), 713-720.

Toppino, T. C. & Bloom, L. C. (2002). The spacing effect, free recall, and two-process theory: A closer look. *Journal of Experimental Psychology, 28,* 437-444.

Toppino, T.C., & Gerbier, E. (2014). About practice: Repetition, spacing, and abstraction. In B.H. Ross (Ed.), *The psychology of learning and motivation: Volume 60: The psychology of learning and motivation* (pp. 113-189). San Diego, CA: Elsevier Academic Press.

Verkoeijen, P. P. J. L., Rikers, R. M. J. P., Schmidt, H. G. (2005). Limitations to the spacing effect. Demonstration of an inverted u-shaped relationship between interrrepetition spacing and free recall. *Experimental Psychology, 52,* 257-263.

Webb, S., & Chang, A. (2012). Vocabulary learning through assisted and unassisted repeated reading. *Canadian Modern Language Review, 68*, 267-290.

Webb, S., & Chang, A. (2015). Second language vocabulary learning through extensive reading with audio support: How do frequency and distribution of occurrence affect learning? *Language Teaching Research*, *19*(6), 667–686.

Tables

Table 1. Descriptive statistics: Vocabulary test scores including means and *SD* in parentheses.

Time	Experimental group		Average
	Intensive (*n* = 33)	Spaced (*n* = 36)	Total (*n* = 69)
Pretest /36	13.97 (4.22)	13.05 (5.29)	13.49 (4.80)
Immediate Posttest /36	25.45 (7.17)	21.66 (6.72)	23.47 (7.15)
Delayed posttest /36	23.69 (7.71)	18.30 (6.80)	20.88 (7.69)

Table 2: Between-groups comparisons of vocabulary gains: Current study vs. Serrano and Huang (2018). Results (statistical significance, effect size). Discrepancies are highlighted.

	Current study	Serrano & Huang (2018)
Immediate gains	Higher for intensive (sig., small)	Higher for intensive (sig., small)
Retention	Higher for intensive (non sig.)	Higher for spaced (sig., large)
Long-term gains	Higher for intensive (sig., large)	Higher for spaced (non sig.)

Sig. = statistically significant.

Table 3: Within-groups comparisons vocabulary learning across testing times: Current study vs. Serrano and Huang (2018). Statistical significance, effect size. Discrepancies are highlighted.

Time	Current study		Serrano & Huang (2018)	
	Intensive	Spaced	Intensive	Spaced
Pre- immediate post	Sig., large	Sig., large	Sig., medium	Sig., small
Immediate post-delayed	Sig., small	Sig., small	Sig., medium	Non sig.
Pre-delayed	Sig., large	Sig., small	Sig., small	Sig., small

Sig. = statistically significant.

Table 4. Vocabulary gains and retention by condition (% in parentheses). Results on incidental learning are from Serrano and Huang (2018).

	Intensive			Spaced		
	Intentional	Incidental	Average	Intentional	Incidental	Average
Pre- immediate post	11.48 (52.11%)	9.02 (38.2%)	9.78 (42.11%)	8.61 (37.51%)	5.65 (25.07%)	7.13 (31.29%)
Immediate post-delayed	-1.76 (-15.33%)	-5.92 (-65.63%)	-3.78 (40.58%)	-3.36 (-39.02%)	-0.47 (-8.32%)	-1.91 (-23.67%)
Pre-delayed	9.72 (44.12%)	3.10 (13.17%)	6.00 (26.02%)	5.25 (22.87%)	5.17 (22.99%)	5.21 (22.93%)

Figure 1. Procedure

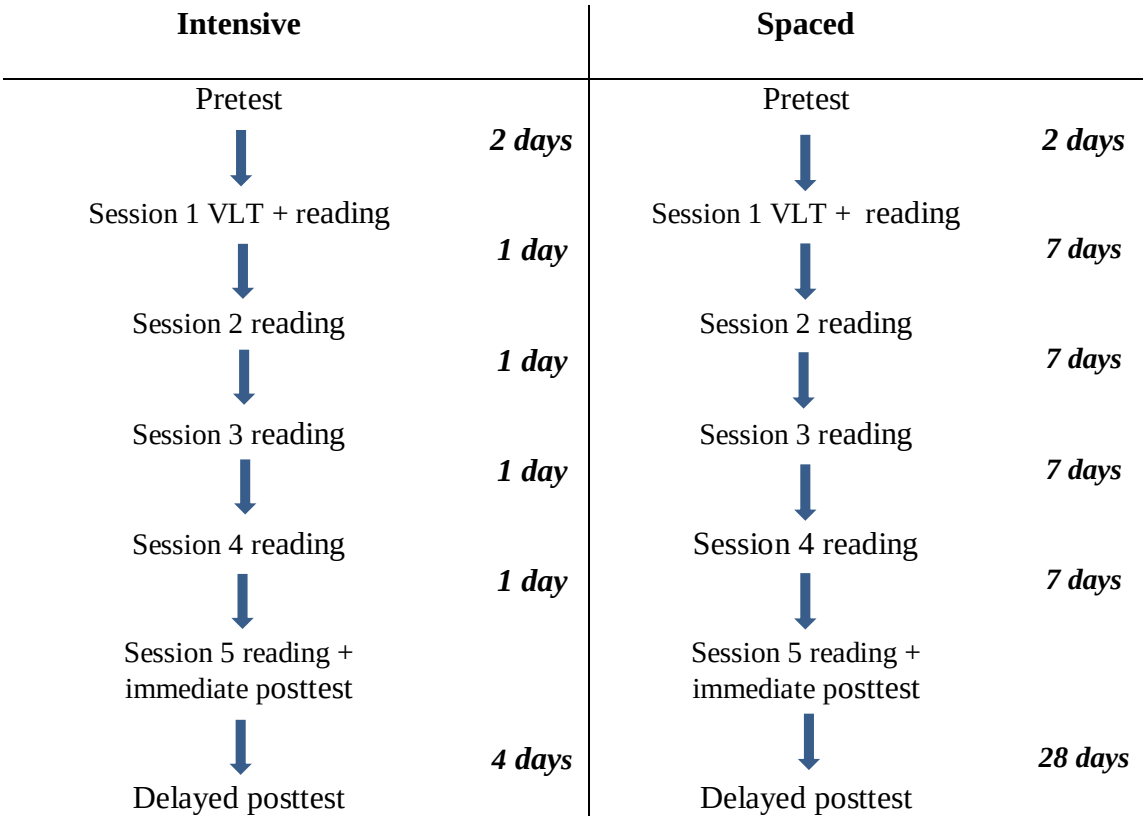
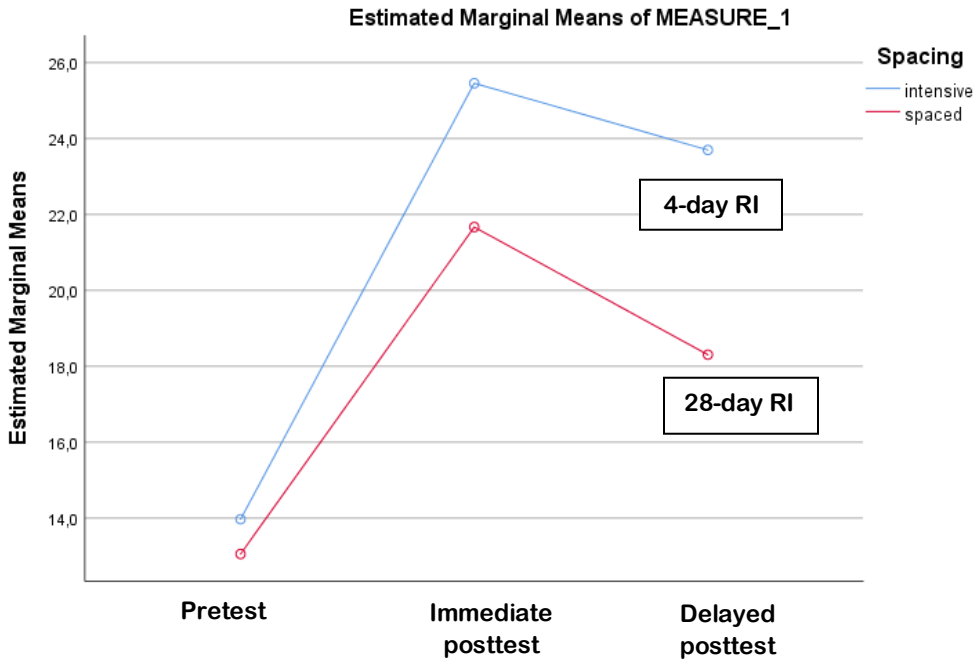


Figure 2: Vocabulary scores through time in the two spacing conditions: intensive vs. spaced



Appendix

Word	Frequency (LexTutor)	Frequency in text
arrangement	1K	1
historic	1K	1
photographed	1K	1
present	1K	1
suggest	1K	1
viewer	1K	1
wood	1K	1
atmosphere	2K	1
attraction	2K	1
capital	2K	1
charm	2K	2
cloth	2K	1
delight	2K	1
district	2K	3
graceful	2K	1
instrument	2K	1
narrow	2K	1
perform	2K	2
skillful	2K	1
spiritual	2K	1
capture	3K	1
literature	3K	1
scatter	3K	1
breeze	4K	1
imitate	4K	2
packet	4K	1
blossom	5K	2
cherry	5K	2
shrine	5K	2
simplicity	5K	1
maple	6K	2
pavilion	6K	2
geisha	11K	4
kimono	11K	2
omamori	-	4
Zen	-	2