Korean Journal of English Language and Linguistics, Vol 22, May 2022, pp. 441-457
DOI: 10.15738/kjell.22..202205.441

KOREAN JOURNAL OF ENGLISH LANGUAGE AND LINGUISTICS

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

http://journal.kasell.or.kr

Effects of Part and Whole Learning on the Learning of L2 Words and Idioms*

Jungyoon Choi · Hyunsook Yoon (Hankuk University of Foreign Studies)

Received: April 10, 2022
Revised: May 15, 2022
Accepted: May 27, 2022

Jungyoon Choi (1st author)
Graduate Student, Dept. of TESOL, Graduate School, Hankuk University of Foreign Studies
Tel: 02) 2173-3978
Email: nadia0308@naver.com

Hyunsook Yoon (corresponding author)
Professor, Dept. of English Education, Hankuk University of Foreign Studies
Tel: 02) 2173-3978
Email: hsyoon3@hufs.ac.kr

* This work was supported by Hankuk University of Foreign Studies Research Fund of 2021.

ABSTRACT

Choi, Jungyoon and Hyunsook Yoon. 2022. Effects of part and whole learning on the learning of L2 words and idioms. *Korean Journal of English Language and Linguistics* 22, 441-457.

This study examined the interrelation between spacing and types of lexical items in SLA. The study compared the effects of four spacing methods (i.e., 1-item part, 5-item part, 10-item part, and 20-item whole learning) on the learning of L2 words and idioms. The experiment was held in a classroom-based environment to suggest the optimum spacing when learning a wordlist and whether the optimum spacing should be applied differently depending on vocabulary types in the real classroom. Eighty adult learners of English in Korea participated in learning L1-L2 paired associates of 20 words and 20 idioms in one of the four spaced conditions with different spacing. The results of the study showed a significant interaction between spacing methods and vocabulary types. On the immediate posttests, whole learning generated better performances than the 1-item and the 5-item part learning when learning words, while all spaced learning performed better than 1-item part learning when learning idioms. On the 1-week delayed posttests, the 10-item part learning produced more correct responses than the 1-item and the 5-item part learning in word learning but outperformed all the other groups in idiom learning. The results of the study suggest that (a) overall, spaced learning has a better effect on learning lists of words and idioms than 1-item part learning, and (b) the optimum spacing should be determined based on retention intervals and the types of vocabulary, including length, difficulty, frequency levels.

KEYWORDS

spacing effect, part learning, whole learning, idioms

1. Introduction

Retention has been an essential issue for vocabulary learning (Schmitt 2010), as forgetting begins soon after learning a new lexical item. Retention is generally improved by repetition. Nation (2013) emphasized that repetition is highly beneficial for learners to recall lexical items and generate longer retention. Repetition studies have been conducted to find effective repetition schedules, including the number of repetitions and the spacing between the repetitions. Repetitions are practiced either massed or spaced (Verkoeijen, Rikers and Schmidt 2005). A well-known notion in repetition studies is *the spacing effect* (Bahrick, Bahrick, Bahrick and Bahrick 1993), which means that spaced repetition produces superior retention effects than massed repetition (Pavlik Jr. and Anderson 2005). Numerous empirical studies have supported the spacing effect (Cull, Shaughnessy and Zechmeister 1996, Karpicke and Roediger 2007, 2010, Logan and Balota 2008, Pavlik Jr. and Anderson 2005, Verkoeijen et al. 2005).

The term *spacing* can be defined in two different ways. First, spacing is the time lag between study episodes, referred to as time-spacing. Second, spacing is the number of items between study episodes, referred to as item-spacing (Cepeda, Pashler, Vul, Wixted and Rohrer 2006). Nakata and Webb (2016) have suggested two distribution types in item-spacing: part learning and whole learning. When vocabulary items on a list are learned as a whole, it is referred to as whole learning, while if they are divided into smaller block sizes, it is referred to as part learning. In a comparison study, the spacing can be relatively short or long. For example, when spacing is comparably longer than the other spacing, it is called long spacing. When spacing is comparably shorter than the other spacing in a study, it is called short spacing. When the length of spacing is between short and long spacing, it is called medium spacing. However, particular spacing can be short spacing in one study, but it can be long spacing in another study. Thus, the meaning of short or long is a relative term since the term is used comparatively within an experiment.

The prior research on the effects of part and whole learning on vocabulary acquisition reported mixed results. According to the *distributed practice effect*, the effectiveness of repetitions depends on the distribution. Items spaced apart during the study are remembered better than those studied all in one go. In other words, longer spacing or whole learning facilitates learning better than shorter spacing or part learning. Many empirical studies have supported the idea (Bahrick and Phelps 1987, Brown 1924, Cepeda, Vul, Rohrer, Wixted and Pashler 2008, Cepeda, Coburn, Rohrer, Wixted, Mozer and Pashler 2009, Crothers and Suppes 1967, Karpicke and Bauernschmidt 2011, Kornell 2009, Metcalfe, Kornell and Finn 2009, Pashler, Zarow and Triplett 2003, Pyc and Rawson 2012). In contrast, the *list-length effect* theory supports part learning, which explains that learners' recallability decreases as the number of items on the vocabulary list increases (Gillund and Shiffrin 1984). On the other hand, Nakata and Webb (2016) found no significant difference between part learning and whole learning. Their research is one of the critical studies in SLA that compared part and whole learning because most spaced learning studies experimented in psychology before their research.

Another important issue in spacing in SLA is the types of lexical items. The lexical items vary from single words to multi-word units. Countless studies have investigated the effects of spaced learning with single-word pairs, but the question about different types of vocabulary learning still exists. Notably, multi-word units have recently gained much attention in the field of SLA. Knowledge of multi-word units is often described and compared with native-like fluency or competence (Schmitt 2010). Nevertheless, the learning aspect of multi-word units is comparatively less researched (Snoder 2017). Among the various multi-word units, idioms are the most neglected category in language learning and use (Maisa and Karunakaran 2013). Therefore, comparing the results of part and whole learning on words and idioms might indicate whether the spacing generating the best performance in learning words can also derive the same result for idiom learning. The present study investigates the effects of part

and whole learning on learning lists of words and idioms. By examining if the effects of various spacing lengths differ depending on the type of vocabulary, this study aims to suggest the optimum interval of item-spacing when learning words and idioms for short-term and long-term retention.

2. Literature Review

2.1 Effects of Part vs. Whole Learning on L2 Vocabulary Learning

This literature review mainly focuses on item-spacing in SLA rather than time-spacing due to the nature of the experiment. Time-spaced learning can be an effective method for vocabulary learning, but it is difficult or impractical for a teacher to spare time to review the learned wordlist unless encountered naturally in the learning materials afterward. In this sense, item-spacing can be a more relevant approach in the real classroom.

In practice, items are learned and repeated with spacing. The critical question is which particular spacing produces the best short-term and long-term retention results. The theoretical perspectives of the distributed practice effect, which includes the spacing effect and the lag effect, have become the mainstay of supporting the advantages of whole learning. As noted earlier, the spacing effect refers to a phenomenon where spaced learning yields superior retention to massed learning (Verkoeijen et al. 2005), and the lag effect explains that learning with long spacing promotes better learning than that with short spacing (Bahrick et al. 1993). Since the late 1800s, there have been considerable research efforts to find the benefits of distributed practice in the field of psychology (Cepeda et al. 2006, Ebbinghaus 1885).

Numerous empirical studies have supported the theoretical perspectives and proved the benefits of whole learning compared to part learning (Cepeda et al. 2006, Karpicke and Roediger 2010, Kornell 2009, Rea and Modigliani 1988, Pyc and Rawson 2007). Cepeda et al. (2006) summarized that longer spacing has a better effect, especially for verbal memory tasks. Karpicke and Roediger (2010) also suggested whole learning with the straightforward idea of maximizing the spacing between study practices regarding how spacing should be distributed. The study of Pyc and Rawson (2007) corroborated the superiority of whole learning over part learning. They researched distributed retrieval practice with restudying and feedback. The participants were 161 university students asked to study 24 Swahili-English word pairs. The study found that the whole learning group recalled more correct responses than the part learning groups. They suggested that the key reason why the group with whole learning performed better than the other groups is that the larger spacing affected students' memory more than the smaller spacing.

Nonetheless, learners have been shown to believe that learning with shorter spacing is more effective than learning with longer spacing (Kornell 2009). However, Kornell (2009) claimed that learning smaller stacks of cards may be favorable, but it was "detrimental" to learning. In his study, 20 university students participated for course credit. There were two conditions: spaced conditions involving learning a large stack of 20 flashcards (i.e., whole learning) and four smaller stacks (i.e., part learning). Students studied 20 digital flashcards, and each card had two synonyms of GRE-type words. The whole learning group studied the whole 20-word pairs (one stack) four times consecutively in the same order. The part learning group studied five-word pairs (four stacks) four times consecutively in the same order before moving to the second stack. After studying, the students took a recall test. The study results revealed that the percentage of correctly recalled items was significantly higher for the session with the stack of 20 cards than that with the smaller stack condition, which supported that learning word pairs with longer spacing enhances memory better than shorter learning. Past studies found that whole learning facilitated

learning, enhanced performance, and generated better results than part learning.

In contrast to the previous studies that supported the advantages of whole learning, other studies argued for the benefits of part learning. Gillund and Shiffrin (1984) proposed a theoretical background based on the *list-length effect,* described as a phenomenon in which the probability of learners' recall or recognition of an item from a vocabulary list decreases as the length of the list increases. For example, when recalling item A in a 10-item word list, nine items become distractors; and when recalling item A in a 20-item word list, 19 items become distractors. Consequently, the distributed target and distractor items overlap more when the list is long. As a result, the recognition performance decreases (Gillund and Shiffrin 1984), which lends support to learning with part learning over whole learning.

When the spacing among learning practices or the learning items is too broad, there tends to be gained detrimental effects because learners cannot retrieve or recall the first presentation (Serrano, Stengers and Housen 2015). In this respect, part learning is a common method of learning vocabulary applied in previous studies. For example, Webb (2005) examined ten target words, and one item was encountered three times before learning the next item in the study. The study of Barcroft and Rott (2010) investigated 24 target words divided into eight items and repeated twice. The study suggested a negative relationship between word learning and word length.

On the other hand, Nakata and Webb (2016) experimented with using a computer program to find if whole learning is more effective than part learning for learning L2 vocabulary. The 91 freshmen at a Japanese university practiced a vocabulary list of the target items with 20 low-frequency English words and three filler words by 4, 10, and 20 items. Each vocabulary item was practiced five times in all groups. The first practice included presenting the target item in L2 and its equivalent L1 for eight seconds. In the second and the third practices, learners were asked to recall the target word's L1 meaning as receptive retrieval, and the fourth and fifth practices involved L2 recall as productive retrieval. The participants took immediate posttests and a 1-week delayed posttest. The posttest scores showed that the whole learning group did not perform better than the part learning group, which was not consistent with previous studies that supported whole learning over part learning (Brown 1924, Crothers and Suppes 1967, Kornell 2009).

To sum up, numerous spacing studies have been conducted to find the optimum schedule for spaced learning. However, their findings are inconsistent, and more importantly, most of the spacing studies on the learning of L2 vocabulary were laboratory-based experiments (Nakata 2015, Nakata and Suzuki 2019, Nakata and Webb 2016), which remains a question of whether the results can be generalized in the mainstream of L2 pedagogy (Snoder 2017). To our best knowledge, only a few studies to date (Küpper-Tetzel, Erdfelder and Dickhäuser 2014, Rogers and Cheung 2020, Serrano and Huang 2018, Serrano, Stengers and Housen 2014, Snoder 2017) were conducted in classroom settings, but most of them investigated L2 vocabulary learning with time-spacing, not item-spacing. Another noticeable feature of previous vocabulary studies with item-spacing in SLA is that they all examined single words. Only two studies adopted multi-word units but with time-spacing (Serrano, Stengers and Housen 2014, Snoder 2017), and none of the studies dealt with multi-word units with item-spacing as target words.

2.2 Effects of Spacing on the Learning of L2 Multi-word Units and the Definition of Idioms

The lexical items longer than single words, including idioms, phrasal verbs, and formulaic language, are called multi-word units, which are far less researched in spaced learning (Snoder 2017). Nattinger and DeCarrico (1992) considered idioms the most crucial lexical category in formulaic language. Mastering idioms is often indicative of native-like fluency (Schmitt 2000). As many L2 learners desire to learn as many idioms as possible, including idioms in the L2 classroom is warranted (Simpson and Mendis 2003). Unlike the emphasis on idioms, however,

L2 students lack awareness of using them in their language production, probably due to insufficient knowledge and processing difficulties. Consequently, language researchers' and teachers' responsibility is to find ways to teach idioms to students more effectively and enable them to use idioms in their language production.

Schmitt (2000) suggested that multi-word units can be mastered only after repeated exposures. Wood (2006) also asserted that repetition practices are fundamental to acquiring idioms and enhancing language competency and proficiency. Furthermore, Choi and Yoon (2018) showed that only 52.9% of multi-word units were learned after the first repetition practice, and thus further repetition practices are recommended and should be planned in the classroom environment to achieve a satisfactory level of learning for multi-word units. Therefore, it stands to reason that repeated practice is likewise necessary for multi-word units. However, despite the importance of repetition practice with multi-word units, most studies on spaced learning have been applied only to learning single-word items in SLA studies.

Snoder (2017) is the only spacing study of L2 vocabulary that investigated multi-word units, verb-noun collocations. Snoder conducted a classroom-based experiment to examine the effects of three vocabulary teaching constructs: involvement load, intentionality, and spacing. "Involvement load" hypothesis (Laufer and Hulstijn, 2001) predicts that word retention is more effective if the motivational-cognitive load is higher when doing a vocabulary task. "Intentionality" includes word processing modes: one is intentional learning, where a posttest is announced to learners, and the other is incidental learning, where the posttest is unannounced to learners. Fifty-nine Swedish learners of English studied English verb-noun collocations. They learned the items three times and took three productive knowledge posttests. The results showed that only intentional learning showed significant effects on learning gains, while spacing, as well as involvement load, did not significantly affect gain scores.

While Snoder's (2017) study is meaningful in that it applied spacing to the learning of multi-word units, spacing in the study was defined by time, and no spacing studies of L2 vocabulary learning have investigated the learning of idioms under item-spacing. Therefore, the present study suggests that it is worth investigating the effects of item-spacing on the learning of L2 words and comparing those with the learning of L2 idioms to examine whether the effects depend on the type of vocabulary.

Earlier scholars (Alexander 1984, Fernando and Flavell 1981) have submitted that the idiom categories include figurative idioms or metaphorical idioms. Other scholars (Drew and Holt 1998, Nunberg, Sag and Wasow 1994) refer to idioms as metaphorical mappings or figures of speech. Sinclair (1991) defined idioms as "a group of two or more words which are chosen together in order to produce a specific meaning or effect in speech or writing" (p.172). The meanings of idioms are not possibly deduced by the definitions or literal meanings of words, but they refer to figurative meanings which can be learned or known by their common usages. Fernando (1996) distinguished idioms from habitual collocations and narrowed the scope. According to her, idioms are "only those expressions which become conventionally fixed in a specific order and lexical form, or have only a restricted set of variants, acquire the status of idioms" (p. 31). She set the categories by distinguishing idioms from non-idiom categories. Also, she recognized the complexity and difficulty of idioms and classified them into different scales. Liu (2003) adopted her definition of idioms which involves three categories: pure (non-literal), semi-literal, and literal. Pure idioms are non-literal and conventional; for example, *pull someone's leg*, and *kick the bucket.* Semi-literal idioms include one or more literal constituents and non-literal subsenses such as *fat chance* and *use something as a stepping stone*. Literal idioms are either invariant or restricted variant; for instance, *according to* and *throw away.* These categories have clarified the definition of idioms and expanded the boundaries to include a variety of multi-word units compared to other past studies. Given that, the present study adopted the definition of Fernando (1996) and Liu (2003).

3. Method

3.1 Research Questions

The spacing in this study was set by adjusting the spacing used in Nakata and Webb (2016), and the one for the idioms was designed using the same as the spacing for learning words to compare the effects of part and whole learning on the different types of vocabulary items. This study aimed to redress the limitations found in Nakata and Webb's (2016) study and broaden the horizons of spacing studies in SLA. First of all, the study compared 4-item part learning, 10-item part learning, and 20-item whole learning and did not compare the smallest number of items, which is 1-item part learning, as opposed to 20-item whole learning. Also, 10-item, which is half of 20-item, was used; thus, 5-item, which is half of 10-item, would seemingly have been more appropriate to be used as shorter spacing instead of 4-item. That makes the spacing of 1-item part (no spaced), 5-item part (short spaced), 10-item part (medium spaced), and 20-item whole (long spaced) learning a more reasonable option for comparison, which was finally adopted and investigated in the present study to examine the learning of 20 single words and 20 idioms. Second, the same number of spacing was used to learn idioms to compare the similarities and differences in the effects of part and whole learning of single and multi-unit words.

The research questions (RQs) that guided this study are as follows:

RQ1: Are there significant differences among 1-item part, 5-item part, 10-item part, and 20-item whole learning in the word and idiom learning conditions on the *immediate* posttest?

RQ2: Are there significant differences among 1-item part, 5-item part, 10-item part, and 20-item whole learning in the word and idiom learning conditions on the *delayed* posttest?

3.2 Participants

Eighty participants were enrolled in English conversation classes for two months. They consisted of 44 male and 37 female Korean learners of English who work at a company in Korea aged 28-39. Their majors were Computer Science, Business, and Industrial Design. They were randomly assigned into four groups: 1-item part, 5-item part, 10-item part, and 20-item whole learning. Initially, all four groups consisted of 20 students each, but one participant in the 1-item part group dropped the class, and one was added to the 5-item part learning class later, leaving still 80 participants' data in total for the analysis. They had similar English education backgrounds and English proficiency levels. All participants had studied English for six years throughout secondary school and one to two years at the college level. Their TOEIC scores ranged between 550 and 650, and their official OPIc levels were intermediate-mid which showed homogeneity among the participants. Also, they took the Vocabulary Size Test (Nation and Beglar 2007), and the average score was 31.52 (SD = 6.12) out of 60 on the frequency levels of the first 1,000-word to the sixth 1,000-word. There were no significant differences in the Vocabulary Size Test scores among the four groups, $F(3, 76) = 0.20$, $p = .842$.

3.3 Materials

The target items were composed of two types that included 20 words and 20 idioms. The 20 words, which consist of 5 letters each, were extracted from the mid-frequency vocabulary: *brisk, spite*, *pluck*, *batch*, *poise*, *chuck*, *blunt*,

sober, *cramp*, *daunt*, *slack*, *lapse*, *peril*, *stark*, *grove*, *perch*, *bleak*, *smear*, *verge*, and *bribe*. All of them are chosen from the 5th 1,000 level of Nation's BNC/COCA lists because they should be unfamiliar items to the participants (Nakata and Webb 2016), but at the same time, necessary items to teach in a real classroom environment. Another reason is that the importance of teaching and learning the mid-frequency vocabulary is often neglected in L2 vocabulary pedagogy even though their benefits have been confirmed in many studies (Schmitt and Schmitt 2014).

The 20 idioms were extracted from the research of Liu (2003) based on the most frequently used idioms: *bite the bullet*, *chew the fat*, *weather the storm*, *burn your boats*, *drop the ball*, *clear the decks*, *try your luck*, *sweeten the pill*, *foot the bill*, *hit the roof, swallow your pride, draw a blank*, *know the ropes*, *pull your leg*, *spill the beans*, *waste your breath*, *hold your horses*, *keep your head*, *push your luck*, and *rock the boat*. Liu (2003) used Corpus of Spoken, Professional American English, Michigan Corpus of Academic Spoken English, and Spoken American Media English to search the identified 9,683 idioms. The results were combined to develop the most frequent idioms' lists and uncover their usage patterns by comparing the three corpora. All of the 20 target idioms, which involve pure and semi-literal idioms, consist of three-word combinations (*verb + article/possessive pronoun + noun*) to maintain the length of each idiom.

All the participants were required to learn the 20 words and 20 idioms as a list depending on the learning schedules. On the left side of the list, 20 English vocabulary items are presented, while the equivalent Korean meaning is provided on the right side of the list (see Appendix). That is, the target words were studied as paired associates where students learned the form of English vocabulary with the meaning of L1 translation (e.g., *brisk* – *활발한*, *know the ropes* –*요령을 잘 알다*). The form of paired associates was chosen for the following reasons. First, learning with paired associates is found to be an effective method for learning L2 vocabulary (Elgort 2011, Nation 2013, Webb 2009). Second, most previous studies on retrieval and distributed practice employed the learning of paired associates (Bahrick and Phelps 1987, Cepeda et al. 2008, 2009, Karpicke and Bauernschmidt 2011, Metcalfe et al. 2009, Pashler et al. 2003, Pyc and Rawson 2012, Nakata and Webb 2016). Recent studies on spacing for vocabulary learning have typically used paired associates (Kornell 2009, Nakata and Webb 2016). Thus, using the same type of materials can make the comparison with prior studies more reasonable.

3.4 Procedure

The experiment was conducted during regular class hours. It involved two rounds of a pretest, learning session, immediate posttest, and delayed posttest, which was held over 36 days. One round was for the word learning session for 14 days, and the other was for idiom learning for the next 14 days. There was an interval of 7 days between the word and the idiom learning sessions. Each learning session lasted 20 minutes during a 1-hour English class. A week before the first learning session, a pretest was administered. A delayed posttest was held a week after the last learning session. Table 1 presents the experimental procedure.

Table 1. Experimental Procedure

Day	Task
Day 1	Pretest for words
Day 8	Learning words + Immediate posttest
Day 15	1-week delayed posttest for words
Day 22	Pretest for idioms
Day 29	Learning idioms + Immediate posttest
Day 36	1-week delayed posttest for idioms

The pretest used a form of a productive recall: providing Korean meaning while asking for writing the equivalent L2 word (e.g., *활발한 – b____*; *요령을 잘 알다 – k_____ ___ _____*). The test included 60 words and 60 idioms, and 20 of each type were selected from the pretest results. The items that students did not know were chosen to remove variables in the students' prior knowledge of target items.

During the learning sessions, instruction involved three stages. The first stage was a presentation, in which the target words/idioms were presented on the PowerPoint slides and shown to students. The second stage was practice. The teacher read the presented English word or idiom, and students repeated it after the teacher. Then the teacher read the equivalent Korean translation, and students read Korean aloud. The third stage was a productive recall. When the teacher presented a Korean word on a slide, the students verbally gave the equivalent English word or idiom and checked the correct answers presented on the slide afterward. The retrieval was performed three times consecutively. The time for the first exposure of a pair of Korean and English words or idioms was 20 seconds, and the total time for learning 20 words/idioms was 20 minutes for presentation, practice, and recall.

While the total time for learning was the same for all participants, they learned the lexical items in one of four spaced conditions with different item-spacing: 1-item part, 5-item part, 10-item part, and 20-item whole learning.

In the 1-item part learning condition, the first item, Item 1, was presented, practiced, recalled three times, and then moved on to the second item, following the same process for the rest of the items. In the 5-item part learning condition, Item 1 to Item 5 were presented, practiced, recalled three times, and then moved to the next 5 items following the same process until the last set of 5 items. The 10-item part learning condition followed the same procedure as the 5-item part learning, except that the block size was 10 instead of 5. In the 20-item whole learning method, Item 1 to Item 20 were presented, practiced, and recalled three times.

Two identical posttests were administered during the study: an immediate posttest and a delayed posttest. Both of the posttests were unannounced to the participants. An immediate posttest was held right after the learning session. One delayed posttest was executed one week after the learning session to assess the retention of learning. The posttests consisted of 20 questions for words and 20 questions for idioms. On the test, Korean words were presented on the left side of the paper test, and students were required to write equivalent English words in the blank next to the Korean words (e.g., *활발한 – b_____*) and idioms (e.g., *요령을 잘 알다 – k_____ ___ _____*), which matched the format of the learning practice. The order of items was randomly changed on the three pre-, immediate, and delayed tests.

3.5 Scoring and Data Analysis

The total score for the immediate and delayed posttest was 20 points for both word and idiom learning conditions. The posttests were assessed in two different levels of sensitivity following Nakata and Webb's (2016) study. The first is strict scoring, in which only correct spelling for a word and an idiom was judged as a point. The second is sensitive scoring, in which one or two spelling mistakes for a single-word item and an idiom were allowed and given a point. Both scoring levels were analyzed, and there were no significant differences between strict and sensitive scoring results. Thus, the present study used sensitive scoring to present the data because it showed the results more precisely. SPSS statistical programs were utilized to analyze the scored data. A mixed-design 2 (vocabulary type: word or idiom) x 4 (group: 1-item, 5-item, 10-item, 20-item) repeated-measure ANOVA was performed to analyze the interaction between groups and vocabulary types on the immediate and delayed posttests. In addition, two-way 4 (item: 1, 5, 10, 20) x 2 (vocabulary type: word, idiom) ANOVAs were performed to reveal the effects of spacing on the tests depending on the conditions of word and idiom learning with p value set at $p < .001$.

4. Results

4.1 Immediate Posttest Performance

The first research question seeks to find if there are significant differences among 1-item part, 5-item part, 10-item part, and 20-item whole learning, and between the learning of single-word items and idioms on the immediate posttest. Table 2 summarizes the average scores on the immediate posttest.

Table 2. Average Scores on the Immediate Posttest

	Word		Idiom	
Group	Mean	*SD*	Mean	*SD*
1-item part learning (*n* = 19)	7.47	3.24	5.00	3.54
5-item part learning (*n* = 21)	8.62	6.29	12.90	4.21
10-item part learning (*n* = 20)	11.00	4.21	11.80	5.33
20-item whole learning (*n* = 20)	14.85	5.67	14.95	5.89

Note. *SD* stands for Standard Deviation. The maximum score is 20.

The descriptive statistics of the immediate posttest showed substantial learning of the students in both word and idiom learning. In order to examine the interaction between vocabulary type and spacing on the immediate posttest, a mixed-design 2 (vocabulary type: word or idiom) x 4 (learning method: 1-item, 5-item, 10-item, 20-item) repeated-measure analysis of variance (ANOVA) was performed with spacing a between-subjects factor and vocabulary type a within-subjects factor. Table 3 presents the results of the repeated-measure ANOVA.

Table 3. Results of Repeated-measure ANOVA for the Immediate Posttest

	Type III Sum of Squares	*df*	*F*	*Sig.*	η_p^2
Vocabulary type x Group	234.232	3	9.214	< .001	.267
Vocabulary type	18.365	1	2.167	.145	.208
Group	1477.056	3	12.761	< .001	.335

The results showed a significant interaction between vocabulary types and learning methods, $F(3, 234.232) = 9.214$, $p < .001$, $\eta_p^2 = .267$, which indicated that statistical differences were found between vocabulary types depending on the groups. The results also found a significant main effect of the group, meaning that there were significant differences among the four groups $F(3, 1477.06) = 12.761$, $p < .001$, $\eta_p^2 = .335$.

In order to determine whether there was a significant difference in the mean gains among the four groups in both word and idiom learning conditions on the immediate posttest, a two-way 2 (vocabulary type: word, idiom) x 4 (learning method: 1-item, 5-item, 10-item, 20-item) ANOVA was performed with learning method a between-subjects factor and vocabulary type a within-subjects factor. The results of ANOVA detected a significant difference among the four groups in both the word learning condition $F(3, 631.671) = 8.309$, $p < .001$, $\eta_p^2 = .247$ and idiom learning condition $F(3, 1079.528) = 16.575$, $p < .001$, $\eta_p^2 = .396$, producing large effect sizes. The results indicated that the differences in correct responses among the 1-item part, 5-item part, 10-item part, and 20-item whole learning in both word and idiom learning conditions were statistically significant on the immediate posttest. A posthoc test of the Bonferroni method of multiple comparisons was also performed for both word and idiom learning conditions to determine which method showed the best performance among the four groups. Table 4

provides multiple comparisons of the four groups for the word and idiom learning condition.

Table 4. Results of Bonferroni Multiple Comparisons for the Word and Idiom Learning Conditions for the Immediate Posttest

		Word			Idiom		
(I) groups	(J) groups	Mean Difference (I-J)	Std. Error	Sig.	Mean Difference (I-J)	Std. Error	Sig.
20-item	1-item	7.380	1.613	< .001	9.950	1.493	< .001
	5-item	6.230	1.573	.001	2.045	1.456	.985
	10-item	3.850	1.592	.108	3.150	1.473	.214
10-item	1-item	3.526	1.613	.191	6.800	1.493	< .001
	5-item	2.381	1.573	.806	-1.105	1.456	1.000
5-item	1-item	1.145	1.594	1.000	7.900	1.475	< .001

In terms of the word learning condition, the multiple comparisons revealed that the 20-item whole learning group outperformed the 1-item part ($p < .001$, $d = 7.380$) and 5-item part learning groups ($p < .001$, $d = 6.230$). The results indicated that the 20-item whole learning is better than the part learning methods when learning a list of words. Regarding the idiom learning condition, the results of the multiple comparisons showed that the 5-item part learning ($p < .001$, $d = 7.900$), 10-item part learning ($p < .001$, $d = 6.800$), and 20-item whole learning ($p < .001$, $d = 9.950$) groups produced a statistically larger number of correct retrievals than the 1-item part learning group. No statistical difference was found among the 5-item part, 10-item part, and 20-item whole learning groups. The results demonstrated that all three spaced learning groups outperformed the 1-item part learning group in idiom learning.

4.2 Delayed Posttest Performance

The second research question asked if there were significant differences among 1-item part, 5-item part, 10-item part, 20-item whole learning, and between the learning of single-word items and idioms on the delayed posttest. Table 5 summarizes the average scores on the delayed posttest.

Table 5. Average Scores on the Delayed Posttest

	Word		Idiom	
Group	Mean	*SD*	Mean	*SD*
1-item part learning ($n = 19$)	0.84	1.07	0.26	0.45
5-item part learning ($n = 21$)	0.24	0.44	0.48	0.87
10-item part learning ($n = 20$)	3.40	3.87	3.00	2.34
20-item whole learning ($n = 20$)	2.05	1.96	1.50	2.37

Note. SD stands for Standard Deviation. The maximum score is 20.

Overall, the delayed posttest results showed a significant drop in the students' correct responses compared to the immediate posttest, indicating relatively low retention of their learning. In order to examine the interaction between vocabulary types and groups on the delayed posttests, a mixed-design 2 (vocabulary type: word or idiom) x 4 (learning method: 1-item, 5-item, 10-item, 20-item) repeated-measure ANOVA was performed with groups a between-subjects factor and vocabulary type a within-subjects factor. Table 6 presents the results of the analysis.

Table 6. Results of Repeated-measure ANOVA for the Delayed posttest

	Type III Sum of Squares	*df*	*F*	*Sig.*	η_p^2
Vocabulary type x Group	4.498	3	.699	.558	.027
Vocabulary type	18.365	1	1.939	.168	.025
Group	448.913	3	11.706	< .001	.316

The results showed no significant interaction between vocabulary type and groups $F(3, 4.498) = .699$, $p = .558$, $\eta_p^2 = .027$, indicating that the groups did not perform differently by vocabulary type. The results also revealed that there was no significant difference between vocabulary type $F(1, 18.365) = 1.939$, $p = .168$, $\eta_p^2 = .025$, but significant differences were detected among the four groups of 1-item part, 5-item part, 10-item part, and 20-item whole learning $F(3, 448.913) = 11.706$, $p < .001$, $\eta_p^2 = .316$.

In order to find whether there was a significant difference in the mean gains among the four groups in word and idiom learning conditions on the delayed posttest, a two-way 2 (vocabulary type: idiom, word) x 4 (learning method: 1, 5, 10, 20) ANOVA was performed with learning method a between-subjects factor and vocabulary type a within-subjects factor. The analysis revealed that there were significant differences in both word $F(3, 118.664) = 7.868$, $p < .001$, $\eta_p^2 = .237$ and idiom learning conditions $F(3, 93.265) = 10.276$, $p < .001$, $\eta_p^2 = .289$, producing large effect sizes. The results of the word and idiom learning conditions indicated that the numbers of correct retrievals were significantly different among the four groups of 1-item part, 5-item part, 10-item part, and 20 whole learning. Subsequently, to determine which method showed the best performance among the four groups, a posthoc test of the Bonferroni method of multiple comparisons was performed for both word and idiom learning conditions. Table 7 presents the results for the word and idiom learning conditions.

Table 7. Results of Bonferroni Multiple Comparisons for the Word and Idiom Learning Conditions for the Delayed Posttest

		Word			Idiom		
(I) groups	(J) groups	Mean Difference (I-J)	Std. Error	Sig.	Mean Difference (I-J)	Std. Error	Sig.
20-item	1-item	1.208	0.718	.581	1.237	0.557	.177
	5-item	1.812	0.701	.070	1.024	0.543	.380
	10-item	-1.350	0.709	.364	-1.500	0.550	.048
10-item	1-item	2.560	0.718	.004	2.740	0.557	< .001
	5-item	3.160	0.701	< .001	2.520	0.543	< .001
5-item	1-item	-0.604	0.710	1.000	0.213	0.551	1.000

The results for the word learning group showed that the 10-item part learning group significantly outperformed the two other groups, which were the 1-item part ($p < .001$, $d = 2.560$) and 5-item part learning ($p < .001$, $d = 3.160$). However, the difference between the 10-item part learning group and the 20-item whole learning group was not statistically significant ($p < .001$, $d = 1.350$). As for the idiom learning condition, the results showed a significant difference only between the 10-item part learning and the other groups, which were the 1-item part ($p < .001$, $d = 2.740$), 5-item part learning ($p < .001$, $d = 2.520$), and 20-item whole learning group ($p < .001$, $d = 1.50$). The results demonstrate that the 10-item part learning group significantly outperformed the other three groups.

5. Discussion

A central finding of the study is that different types of spacing intervals interacted with lexical items. The immediate posttest found a significant interaction between spacing and vocabulary type. On the immediate posttest, the word learning results showed that the 20-item whole learning group generated significantly better performance than the 1-item and 5-item part learning groups *(20-item > 1-item, 5-item)*. Overall, the results are consistent with numerous previous studies that found the superiority of learning with longer over shorter spacing (e.g., Cepeda et al. 2006, Karpicke and Roediger 2010, Kornell 2009). On the other hand, the results are incongruent with Nakata and Webb's (2016) finding that the whole learning group did not perform better than the part learning group. The difference may result from the fact that the number of spacing in part and whole learning was the same in their study. The two types of learning used 19 items, including three filler items, as spacing. In contrast, the present study used the items themselves as spacing so that the number of spacing in each method was different: 0 for the part, 4 for the 5-item part, 9 for the 10-item part, and 19 for the 20-item whole learning. In this regard, the better results in the 20-item whole learning compared to the 5-item part learning can be explained by the delay of the first retrieval (Karpicke and Roediger 2007). According to prior research by Benjamin, Bjor and Schwartz (1998), Gardiner, Craik and Bleasdale (1973), and Karpicke and Roediger (2007), delaying the initial retrieval increases the difficulty of the initial retrieval, and the increased difficulty provokes much effort, which eventually results in better learning.

In the idiom learning condition, the immediate posttest results revealed that the three spaced learning groups outperformed the 1-item part learning group *(5-item, 10-item, 20-item > 1-item)*. The results indicated that the spacing interval does not significantly affect the posttest results as long as the repetition is not massed. This result is supported by the body of theoretical and empirical research on the spacing effect, which shows the advantage of spaced practices compared to massed practices when learning new materials (Bahrick et al. 1993, Verkoeijen et al. 2005, Nakata and Webb 2016, Snoder 2017). Notably, while the whole learning showed a better performance than part learning in the word learning condition, the idiom learning condition showed no differences among the spaced groups. One possible reason for the results is the characteristics of the vocabulary type. Each idiom consisted of three words, implying that the idiom's length includes spacing since an item counts for spacing in item-spacing research (Cepeda et al. 2006). Therefore, the spacing in each idiom might have taken part in the spacing effect in the 5-item part, 10-item part, and 20-item whole learning.

In contrast to the immediate posttest that revealed a significant interaction between spacing and vocabulary type, the delayed posttest found no interaction between them but only a significant main effect of spacing. It can be instantly interpreted that the length of spacing has a greater impact on vocabulary learning than the vocabulary type in the longer term. However, the instant interpretation needs caution, considering the various characteristics of vocabulary items, which will be explicated later. The delayed posttest results in the word learning condition showed that the 10-item part learning group resulted in better performance than the 1-item part and the 5-item part learning *(10-item > 1-item, 5-item)*. In the idiom learning condition, the 10-item part learning group outperformed the other three groups *(10-item > 1-item, 5-item, 20-item)*. The findings that the 10-item part learning showed better learning gains than the 1-item and 5-item part learning in both word and idiom learning are supported by theoretical backgrounds of spacing effect (Bahrick et al. 1993) and distributed practice effect (Gillund and Shiffrin 1984). Interestingly, however, there was a difference in the vocabulary type: the 10-item part learning performed better than the 20-item whole learning in idiom learning, but it did not happen in the word learning condition. According to the *task difficulty framework* (Karpicke and Roediger 2007, Nakata and Webb 2016, Spitzer 1939), item difficulty includes familiarity, imageability, L1 word frequency, or semantic clustering. Those factors cause interference and slow down initial learning. When the item difficulty meets the desirable difficulty requiring a

considerable but desirable amount of effort in a learning task, the slow learning in the initial stage facilitates learning for the long-term, which is called the *desirable difficulty framework* (Bjork 1994). The task difficulty framework also prescribes that reducing spacing enhances retention when the task materials are relatively difficult (Karpicke and Roediger 2007, Nakata and Webb 2016, Spitzer 1939). The items of idiom learning in this study consisted of three words and pure or semi-literal idioms, which can lead to low imageability. In other words, the idiom's length and low imageability might have made the task difficulty high. In this sense, the 20-item whole learning might have been more demanding for the students. Consequently, the 10-item part learning was found to be the optimum spacing when learning those idioms for the long-term effect.

6. Pedagogical Implications

While the efforts to find the benefits of spaced learning have been continued in psychology since the late 1800s (Cepeda et al. 2006, Ebbinghaus 1885), the investigations of applying different types of lexical items have been less researched in the field of SLA. This study is the first kind that compared word and idiom learning with spaced learning. The most crucial pedagogical implication of the study is that different spacing interacts with different lexical items. Spaced learning should be supported by designing vocabulary learning in the syllabus depending on vocabulary types when used to practice in the real classroom.

Naturally, the optimum spacing in one study may not be the best choice of spacing when the target item is not the same or similar, including the length, frequency, and difficulty of items. According to the desirable difficulty framework, increasing difficulty in learning over time is effective (Bjork 1999), and increasing spacing between repeated practices enhances learning when the task difficulty is relatively low (Spitzer 1939). Conversely, decreasing spacing between repeated practices enhances learning when the task difficulty is relatively high. For example, when the target words consist of low-frequency vocabulary, the idioms consist of unfamiliar vocabulary, or the learner's English proficiency level is lower than intermediate, the spacing of repeated practices is better to be reduced to facilitate learning. On the other hand, when the task difficulty is relatively low, the spacing can be increased. For example, when the target words consist of high-frequency vocabulary, the idioms consist of familiar vocabulary, or the English proficiency level of learners is higher than intermediate, the spacing of repeated practices is better to be increased to enhance learning.

7. Concluding Remarks

The present study aimed to determine whether learners' performances show differences between learning words and idioms depending on the length of spacing and suggest the optimum spacing among the different schedules of part and whole learning when learning a list of words and idioms for the short-term and long-term retention. The important finding is that there were significant differences between word and idiom learning depending on spacing.

Although this study provides significant implications for the spaced learning studies in SLA, it is not without limitations. Firstly, more participants would provide better reliability to this study, and the participants' proficiency levels can be examined as another variance. Secondly, in this study, within-session spacing, not between-session spacing, was manipulated so that the application of the results is limited to the type of learning session. Thirdly, the study adopted equal spacing. Different types of spacing (i.e., expanding spacing and contracted spacing) might find different directions for idiom learning since some of the studies showed advantages of expanding spacing

over equal spacing (Cull et al. 1996, Greene 2008, Nakata 2015, Karpicke and Roediger 2007, Landauer and Bjork 1978). Lastly, different types of vocabulary should be investigated as target vocabulary. The present study used only mid-frequency vocabulary for single words and pure and semi-literal items for idioms. Research on different types of vocabulary items might generate different results. For example, previous spaced learning studies showed different results depending on the types of paired associates, such as face-name (e.g., Laundauer and Bjork 1978), L1-L2 (e.g., Karpicke and Roediger 2007), L2-L1 (e.g., Kang, Lindsey, Mozer and Pashler 2014) and cue-target (e.g., Logan and Balota 2008). Therefore, there is a high probability of generating significant differences among the different vocabulary types, which will give a new perspective on spaced learning research.

References

Alexander, R. 1984. Fixed expressions in English: Reference books and the teacher. *ELT Journal* 38(2), 127–132..

Bahrick, H. P., L. E. Barhick, A. S. Bahrick and P. E. Bahrick. 1993. Maintenance of foreign language vocabulary and the spacing effect. *Psychological Science* 4, 316–321.

Bahrick, H. P. and E. Phelps. 1987. Retention of Spanish vocabulary over 8 years. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 13(2), 344–349.

Barcroft, J. and S. Rott. 2010. Partial word form learning in the written mode in L2 German and Spanish. *Applied Linguistics* 31(5), 623–650.

Benjamin, A. S., R. A. Bjork and B. L. Schwartz. 1998. The mismeasure of memory: When retrieval fluency is misleading as a metamnemonic index. *Journal of Experimental Psychology: General* 127(1), 55–68.

Bjork, R. A. 1994. Memory and metamemory considerations in the training of human beings. In J. Metcalfe and A. Shimamura, eds., *Metacognition: Knowing about Knowing,* 185–205. Cambridge, MA: MIT Press.

Bjork, R.A. 1999. Assessing our own competence: Heuristics and illusions. In D. Gopher and A. Koriat, eds., *Attentin and Performance XVII: Cognitive Regulation of Performance: Interaction of Theory and Application,* 435–459. Cambridge, MA: MIT Press.

Brown, W. 1924. Whole and part methods in learning. *Journal of Educational Psychology* 15, 229–233.

Cepeda, N. J., N. Coburn, D. Rohrer, J. T. Wixted, M. C. Mozer and H. Pashler. 2009. Optimizing distributed practice: Theoretical analysis and practical implications. *Experimental Psychology* 56(4), 236–246.

Cepeda, N. J., H. Pashler, E. Vul, J. T. Wixted and D. Rohrer. 2006. Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin* 132(3), 354–380.

Cepeda, N. J., E. Vul, D. Rohrer, J. T. Wixted and H. Pashler. 2008. Spacing effects in learning: A temporal ridgeline of optimal retention. *Psychological Science* 19(11), 1095–1102.

Choi, J. and H. Yoon. 2018. The repetition effects on learning formulaic language in form-focused instruction. *Modern English Education* 19(1), 76–85.

Crothers, E. and P. Suppes. 1967. *Experiments in Second Language Learning*. New York, NY: Academic Press.

Cull, W. L., J. J. Shaughnessy and E. B. Zechmeister. 1996. Expanding understanding of the expanding-pattern-of-retrieval mnemonic: Toward confidence in applicability. *Journal of Experimental Psychology: Applied* 2, 365–378.

Drew, P. and E. Holt. 1988. Complainable matters: the use of idiomatic expressions in making complaints. *Social Problems* 35(4), 398–378.

Ebbinghaus, H. 1885. *Memory: A Contribution to Experimental Psychology.* Republished 1964. New York: Dover.

Elgort, I. 2011. Deliberate learning and vocabulary acquisition in a second language. *Language Learning* 61(2),

367–413.
Fernando, C. 1996. *Idioms and Idiomaticity*. Oxford: Oxford University Press.
Fernando, C. and R. Flavell. 1981. *On Idiom: Critical Views and Perspective* (Vol.5). Exeter: University of Exeter.
Gardiner, F. M., F. I. Craik and F. A. Bleasdale. 1973. Retrieval difficulty and subsequent recall. *Memory & Cognition* 1(3), 213–216.
Gillund, G. and R. M. Shiffrin. 1984. A retrieval model for both recognition and recall. *Psychological Review* 91(1), 1–67.
Greene, R. L. 2008. Repetition and spacing effects. In H. L. Roediger III, ed., *Learning and Memory: A Comprehensive Reference. Vol. 2: Cognitive Psychology of Memory,* 65–78. Oxford: Elsevier.
Kang, S. H., R. V. Lindsey, M. C. Mozer and H. Pashler. 2014. Retrieval practice over the long term: Should spacing be expanding or equal-interval? *Psychonomic Bulletin & Review* 21(6), 1544–1550.
Karpicke, J. D. And A. Bauernschmidt. 2011. Spaced retrieval: Absolute spacing enhances learning regardless of relative spacing. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 37(5), 1250–1257.
Karpicke, J. D. and H. L. Roediger. 2007. Repeated retrieval during learning is the key to long-term retention. *Journal of Memory & Language* 57, 161–162.
Karpicke, J. D. and H. L. Roediger. 2010. Is expanding retrieval a superior method for learning text materials? *Memory & Cognition* 38(1), 116–124.
Kornell, N. 2009. Optimising learning using flashcards: Spacing is more effective than cramming. *Applied Cognitive Psychology* 23(9), 1297–1317.
Küpper-Tetzel, C. E., E. Erdfelder and O. Dickhäuser. 2014. The lag effect in secondary school classrooms: Enhancing students' memory for vocabulary. *Instructional Science* 42(3), 373–388.
Landauer, T. K. and R. A. Bjork. 1978. Optimum rehearsal patterns and name learning. *Practical Aspects of Memory* 1, 625–632.
Nattinger, J. R. and J. S. DeCarrico. 1992. *Lexical Phrases and Language Teaching*. Oxford University Press.
Laufer, B. and J. Hulstijn. 2001. Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics* 22(1), 1–26.
Liu, D. 2003. The most frequently used spoken American English idioms: A corpus analysis and its implications. *TESOL Quarterly* 671–700.
Logan, J. M. And D. A. Balota. 2008. Expanded vs. equal interval spaced retrieval practice: Exploring different schedules of spacing and retention interval in younger and older adults. *Aging, Neuropsychology, and Cognition* 15(3), 257–280.
Maisa, S. and T. Karunakaran. 2013. Idioms and importance of teaching idioms to ESL students: A study on teacher beliefs. *Asian Journal of Humanities and Social Sciences (AJHSS)* 1(1), 110–122.
Metcalfe, J., N. Kornell and B. Finn. 2009. Delayed versus immediate feedback in children's and adults' vocabulary learning. *Memory & Cognition* 37(8), 1077–1087.
Nakata, T. 2015. Effects of expanding and equal spacing on second language vocabulary learning: Does gradually increasing spacing increase vocabulary learning? *Studies in Second Language Acquisition* 37(4), 677–711.
Nakata, T. and Y. Suzuki. 2019. Effects of massing and spacing on the learning of semantically related and unrelated words. *Studies in Second Language Acquisition* 41(2), 287–311.
Nakata, T. and S. Webb. 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(3), 523–552.

Nation, I. S. 2013. *Learning Vocabulary in Another Language* (2nd ed.). Cambridge: Cambridge University Press.
Nunberg, G., I. A. Sag and T. Wasow. 1994. Idioms. *Language* 70(3), 491–538.
Pashler, H., G. Zarow and B. Triplett. 2003. Is temporal spacing of tests helpful even when it inflates error rates? *Journal of Experimental Psychology: Learning, Memory, and Cognition* 29(6), 1051–1057.
Pavlik Jr, P. I. and J. R. Anderson. 2005. Practice and forgetting effects on vocabulary memory: An activation-based model of the spacing effect. *Cognitive Science* 29(4), 559–586.
Pyc, M. A. and K. A. Rawson. 2007. Examining the efficiency of schedules of distributed retrieval practice. *Memory & Cognition* 35(8), 1917–1927.
Pyc, M. A. and K. A. Rawson. 2012. Why is test–restudy practice beneficial for memory? An evaluation of the mediator shift hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 38(3), 737–746.
Schmitt, N. 2000. *Vocabulary in Language Teaching.* Cambridge: Cambidge University Press.
Schmitt, N. 2010. *Researching Vocabulary: A Vocabulary Research Manual*. Springer.
Rea, C. P. and V. Modigliani. 1988. Educational implications of the spacing effect. In M. M. Gruneberg, P. E. Morris and R. N. Sykes, eds., *Practical Aspects of Memory: Current Research and Issues, Vol. 1. Memory in Everyday Life,* 402–406. John Wiley & Sons.
Rogers, J. and A. Cheung. 2020. Does it matter when you review?: Input spacing, ecological validity, and the learning of L2 vocabulary. *Studies in Second Language Acquisition* 1–19.
Serrano, R. and H. Y. Huang. 2018. Learning vocabulary through assisted repeated reading: How much time should there be between repetitions of the same text? *TESOL Quarterly* 52(4), 971–994.
Simpson, R. and D. Mendis. 2003. A corpus-based study of idioms in academic speech. *TESOL Quarterly* 37(3), 419–441.
Schmitt, N. 2010. *Researching Vocabulary: A Vocabulary Research Manual*. Springer.
Schmitt, N. and D. Schmitt. 2014. A reassessment of frequency and vocabulary size in L2 vocabulary teaching. *Language Teaching* 47(4), doi:10.1017/S0261444812000018.
Serrano, R., H. Stengers and A. Housen. 2015. Acquisition of formulaic sequences in intensive and regular EFL programmes. *Language Teaching Research* 19(1), 89–106.
Snoder, P. 2017. Improving English Learners' Productive Collocation Knowledge: The Effects of Involvement Load, Spacing, and Intentionality. *TESL Canada Journal* 34(3), 140–164.
Spitzer, H. F. 1939. Studies in retention. *Journal of Educational Psychology* 30(9), 641–656.
Verkoeijen, P. P., R. M. Rikers and H. G. Schmidt. 2005. Limitations to the spacing effect: Demonstration of an inverted u-shaped relationship between interrepetition spacing and free recall. *Experimental Psychology* 52(4), 257–263.
Webb, S. 2005. Receptive and productive vocabulary learning: The effects of reading and writing on word knowledge. *Studies in Second Language Acquisition* 27(1), 33–52.
Webb, S. 2009. The effects of receptive and productive learning of word pairs on vocabulary knowledge. *RELC Journal* 40(3), 360–376.
Wood, D. 2006. Uses and functions of formulaic sequences in second language speech: An exploration of the foundations of fluency. *Canadian Modern Language Review* 63(1), 13–33.

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

Appendix

Task Materials for Single-Word Items and Idioms

	Words		Idioms	
1	brisk	활발한	bite the bullet	이를 악물고 하다
2	spite	악의	chew the fat	오래 담소를 나누다
3	pluck	뽑다	weather the storm	고비를 넘기다
4	batch	집단	burn your boats	돌이킬 수 없게 만들다
5	poise	침착성	drop the ball	실수로 망치다
6	chuck	던지다	clear the decks	전투 준비를 하다
7	blunt	무딘	try your luck	되든 안되든 해보다
8	sober	술 취하지 않은	sweeten the pill	불쾌한 상황을 포장하다
9	cramp	경련	foot the bill	한턱 내다
10	daunt	기죽게 하다	hit the roof	벌컥 화내다
11	slack	느슨한	swallow your pride	자존심을 버려라
12	lapse	깜빡함	draw a blank	아무 결과를 얻지 못하다
13	peril	심각한 위험	know the ropes	요령을 잘 알다
14	stark	삭막한	pull your leg	농담이야
15	grove	수풀	spill the beans	비밀을 누설하다
16	perch	걸터앉다	waste your breath	말해봐야 입만 아프다
17	bleak	암울한	hold your horses	흥분부터 하지 마라
18	smear	문지르다	keep your head	냉정을 잃지 않다
19	verge	길가	push your luck	자만하다
20	bribe	뇌물	rock the boat	평지풍파를 일으키다