The Effects of Repeated Oral Reading Practice on the Retention of High-Frequency Multiword Items for EFL Learners: Multiple Dimensions

Anna C-S Chang
Hsing Wu University
<annachang@livemail.tw>

Abstract

Research has shown that second language (L2) learners generally lack multiword expression knowledge, and L2 researchers and practitioners have tried various techniques to assist L2 learners to acquire it more efficiently. This study adopted an under-researched technique—repeated oral reading—to enhance the retention of high-frequency multiword items by 62 EFL college students divided into experimental (n =38) and control (n =24) groups. Fifteen unfamiliar multiword items comprising only known individual words were selected through a pre-test based on a theme-based text. All students received a formal instruction first, followed by the experimental group orally reading the text six times under a time constraint. A two-week delayed post-test was used to test students' retention of four dimensions of multiword knowledge: aural forms and aural meanings, and written meanings and use. Except for use, the experimental group significantly outperformed the control group in the other three dimensions. Four fixed factors (oral reading speed, prior vocabulary knowledge, dimensions of multiword knowledge, and the number of words per item) were analyzed via GLMM. Results showed three factors had significant effects on retaining multiword items except oral reading speed. Based on the results, pedagogical implications are discussed and suggestions are made.

Keywords: reading aloud, oral reading, repeated oral reading, multiword items, high-frequency words

Recent research found that L2 learners lack multiword item (MWI) knowledge, and they are familiar with far more high frequency individual words than MWIs that comprise high frequency words (e.g., Kim, 2016; Nguyen & Webb, 2017; Park & Chon, 2019); however, studies have also shown that approximately 50% to 70% of spoken and written discourse consists of fixed MWIs (Erman & Warren, 2000). An MWI is used in this study as an umbrella term for strings of language items containing more than a single word (Siyanova-Chanturia &

Omidian, 2020), which may be seen as "idioms," "collocations," "lexical bundles," or "lexical chunks" in the literature. Many researchers have investigated how to increase learning effectiveness through various instructional methods, such as looking up the target items for meanings in dictionaries (Dziemianko, 2010; Laufer, 2011); deliberately learning unfamiliar collocations through their L1 translations or by doing various exercises (Webb & Kagimoto, 2009, 2011); using sound repetition (Lindstromberg & Boers, 2008); grouping idioms by themes or different verbs (Zyzik, 2011), to name only a few; however, one old-fashioned but commonly used technique—reading aloud (used interchangeably with oral reading), is apparently under-researched or a missing piece in learning MWIs (Alali & Schmitt, 2012).

Oral reading is often considered boring, dated, or anxiety-provoking with learners barking out texts without comprehension (see Gibson, 2008 for the strengths and weaknesses of oral reading), but many psycholinguistic studies have shown the importance of vocalization in learning an L2. According to Baddeley (2013), the auditory memory traces decay very rapidly in the phonological loop of working memory; however, they can revive through articulatory rehearsal and be restored in the long-term memory system. Many autonomous learners also use this technique as part of private study to practice pronunciation and vocabulary, to build up confidence in speaking, and to improve fluency (Gibson, 2008). Moreover, teachers use it as a diagnostic tool to check students' pronunciation problems, graphemic-phonemic connections, etc. Reading aloud, when repeated many times, is called repeated oral reading. Many recent studies have demonstrated the importance of "repetition" or "frequency of encounter" in L2 vocabulary learning (see Uchihara, Webb, & Yanagisawa, 2019), but few studies have examined how repeated oral reading affects MWI learning. Repeated oral reading offers at least three advantages over facilitating learning. Firstly, it is repetitive, which may increase the number of opportunities for encountering the same vocabulary words; secondly, vocalizing the newly learned linguistic elements may help learners to better retain what is learned (Gathercole & Conway, 1988). Thirdly, students enjoy the experience of bringing the language to life through their voice when they practice class content orally (Chang, 2019). For these encouraging reasons, the current study adopted repeated oral reading to examine the extent of EFL (English as a foreign language) students' MWI learning. If repeated oral reading can be demonstrated to improve retention of MWI learning for L2 learners, it will increase the practice varieties for vocabulary learning.

Background

Repeated Oral Reading in L2 Vocabulary Learning

Repeated oral reading has been used extensively in L1 contexts for developing reading fluency and literacy, as well as for remedial purposes to assist struggling readers (Rasinski, Homan, & Biggs, 2009). In L2 contexts, repeated oral reading is normally conducted in chorus (see Alali & Schmitt, 2012) rather than individually; this is because individual oral reading is very time consuming and many language teachers simply cannot afford the time to do it. Though some researchers were able to employ individual oral reading, the number of participants was limited (e.g., Chang, 2019; Lin, 2016), and the main focuses were on improving oral reading rates, comprehension, and pronunciation rather than on vocabulary retention (Shimono, 2019). In the limited literature, three relevant studies into reading aloud on vocabulary learning, are discussed below:

To investigate whether doing oral practice or written practice makes a difference in acquiring vocabulary knowledge, Alali and Schmitt (2012) conducted a study of 35 Kuwaiti female junior high school students' acquisition of single words and idioms. In the study, oral practice involved the students reading target formulaic sequences 10 times in chorus for ten minutes; in the written review, students worked in groups for 10 minutes on a written recall task. Thirty idioms were selected; an unknown word from each idiom made up the target single words. The results showed that oral practice and written recall resulted in a similar pattern of learning of single words and idioms. However, written repetition was consistently more effective than oral reading for form and meaning recognition and recall. Students who did oral practice retained 0.50/10 words for the idiom form recall task and 2.30/10 for the meaning recall in the 12-day delayed post-test. Acquisition rates for idiom recall tasks were not very satisfactory, possibly because oral reading was done as a group; it may be possible to improve learning rate by conducting oral reading on an individual basis. Another significant reason could be that each idiom contained an unknown word, which might have increased the difficulty in acquisition (Kim, 2016; Zyzik, 2011). Could the acquisition rate be enhanced if all items comprised of familiar words? This study hence attempts to fill the two gaps.

Another study by Durrant and Schmitt (2010) examined the effects of learning 20 pairs of words with adjectives and nouns under three different conditions: presenting each collocation in a sentential context only once; orally repeating a collocation two times, and presenting a target collocation in different contexts. Participants were 84 non-native adult speakers of English from various countries. An uninformed cued recall test was given after the treatment. Results showed that the two experimental repetition conditions had an effect superior to the no repetition condition. The verbatim repetition of a single sentential context (Cohen's $d = .56$; $Mdn = 4.5/10$) also revealed better learning outcomes over the exposure to varied contexts (Cohen's $d = .48$; $Mdn = 5.0/10$). The authors concluded that reading identical sentences aloud twice might ease the cognitive load and hence improve collocational memory trace. Additionally, oral reading allowed participants to dedicate more attention to the language.

A more recent study by Chang and Chen (2022) investigated the effects of written and oral vocabulary exercises on EFL learners' retention of single-word items (SWI) and multi-word items (MWI). After receiving formal instruction of four texts, the written exercises group did gap-filling, meaning-matching, and rearranging scrambled sentences, whereas the oral vocabulary exercise group was given a list of the target items embedded in sentences and orally practiced the target items six times. It must be mentioned that students' oral reading speeds were not measured. Sixty unfamiliar vocabulary items were tested in four dimensions of vocabulary knowledge: aural form and meaning recall, L2 written form and meaning recall. Mixed findings were found in the one-week delayed post-tests. The results showed that students performed significantly better on the MWIs than SWIs regardless which types of vocabulary exercises were adopted. In terms of vocabulary dimensions, students performed best for L2 written meaning recall (66%), followed by aural form recall (59%), aural meaning recall (51%), and L2 form recall (50%). According to the authors, all SWIs were unfamiliar to students in all dimensions tested, which made it difficult for them to guess its meaning, but the MWIs were comprised mostly of known individual words, which might have made the tests easier for them to guess meanings from the other words in the MWI.

Taken altogether, although no conclusive findings were found, the studies suggest that repetition may play some role in vocabulary learning, and varying types of repetition also make

some differences in learning outcome. One area that is not clear is whether a learner who reads more fluently can retain more vocabulary knowledge, because oral reading fluency has been used to measure a learner's reading competence or their familiarity with the content (e.g., Fuchs, Fuchs, Hosp, & Jenkins, 2001; Jiang, 2016). Furthermore, the better a learner is, the more fluently they can read with appropriate prosody (Grabe, 2009; Kuhn & Stahl, 2003). Having said this, no L2 research up to the present has investigated whether L2 learners who read a text more fluently can retain more MWIs. Therefore, in the present study, student participants' reading speed was included as a variable to examine its effect on learning multiword items.

Learners' Prior Vocabulary Knowledge and Vocabulary Learning

Many studies that took into account learners' prior vocabulary knowledge have shown that learners with a larger vocabulary size tend to pick up more words from reading (e.g., Horst, Cobb, & Meara, 1998; Webb & Chang, 2015; Zahar, Cobb, & Spada, 2001) or viewing (Peters & Webb, 2018). For example, Horst, Cobb, and Meara (1998) examined the relationship between vocabulary knowledge and incidental learning gains through reading and listening to a graded reader. Students' vocabulary knowledge was measured through the Vocabulary Levels Test (Nation, 1983). A medium-sized relationship was found for the two variables. Similar results were shown in an extensive reading program by Webb and Chang (2015), whose students were divided into high-, intermediate-, and low- levels according to their scores on a pretest measuring knowledge of target vocabulary. They found that the more prior vocabulary knowledge students had, the greater their gains from extensive reading. Measuring students' meaning recall and meaning recognition in their two separate experiments, Peters and Webb (2018) also reported a positive relationship between learners' prior vocabulary knowledge and vocabulary learning through viewing TV.

Although the conclusions drawn above are specific to acquisition of single words, a study by Peters (2016) suggests that these findings can be generalized to MWI. Peters found that the larger a learner's vocabulary size, the higher the odds that they could recall a new collocation; specifically, she found that the odds to recall a new collocation in the first form recall test were 13 times higher for every 100 words more in a learner's vocabulary size. Taken together, these research findings suggest that prior vocabulary knowledge may affect the ability of L2 learners to acquire single word. The reason for this may be that students with greater vocabulary knowledge are likely to have greater text comprehension (Laufer, 2013; Webb & Chang, 2015), and this may allow them to pay greater attention to unknown words.

Number of Words in Multiword Items

The characteristics of MWI may affect the extent to which an MWI can be learned. These characteristics involve the following: whether the items being semantically transparent or opaque (Gyllstad & Wolter, 2016; Macis & Schmitt, 2017), the presence of known or unknown words in an expression (Kim, 2016), the combination (verb-noun, adjective-noun) of collocations (Peters, 2016), number of node words (Webb & Kagimoto, 2011), and thematical groupings (Zyzik, 2011). However, few studies have taken into account how the numbers of words in a MWI affect learning outcome. In general, the longer a single word is, the more difficult it is to learn. Campoy (2008) found that shorter words (four-phonemes) were better recognized than strings of long words (six-phonemes). Peters (2016) also found an MWI that comprises of longer individual words, the lower chance for a correct response. In the same

vein, an MWI containing three or more words is supposed to be more difficult to learn than one consisting of only two words. Hence, this study included this variable to determine how it affected MWI acquisition.

Measuring Different Dimensions of Vocabulary Knowledge

Word knowledge is not an all-or-nothing construct. According to Nation and Webb (2011), knowing a word involves 18 dimensions. Some learning tasks may be more effective in gaining receptive knowledge (e.g., how does the word look like?) whereas others in productive knowledge (e.g., how is the word spelled?). Therefore, it is important to measure different aspects of word knowledge in a test because it "may provide a much more clear evaluation of the relative efficacy of tasks" (Webb, 2005, p. 50). Many recent studies involved several tests to measure students' learning efficacy. For example, Teng (2016) measured his students' gains from reading a graded reader in four dimensions: form, grammar, meaning, and collocation. Pigada and Schmitt (2006) measured a learner's spelling, meaning, and grammatical characteristics gains from reading four texts. The two studies reported that their students gained significantly more on spelling than on grammar. In a long-term narrow reading study, Chang (2019), however, found that L2 students made significantly more improvement in their knowledge of written and aural meanings than for spelling and use. In general, requiring students to produce answers is always more difficult than having them to select one from a few options, but to determine which aspects of knowledge should be measured must depend on the learning conditions and tasks.

Based on the limited number of studies reviewed above, the effects of repeated oral reading on L2 learners' vocabulary acquisition are mixed (Alali & Schmitt, 2012; Chang & Chen, 2022; Durant & Schmitt, 2010). The effects of reading aloud in chorus may be different from the effects of reading aloud individually. This study therefore adopted individual repeated oral reading. Fifteen unknown MWIs comprising only known individual words were selected from a theme-based story related to "time." Four dimensions of vocabulary knowledge were assessed: spelling, aural meaning, written meaning, and use. Two research questions were addressed. The first question focused on comparing the difference in retention for four dimensions of multiword items between the experimental group and the control group. The second research question explored mainly the learning effectiveness for the experimental group in terms of their vocabulary knowledge, vocabulary dimensions, reading rates, and the number of words in each item. The two research questions are as below:

RQ1: Were there any significant differences in the retention for four dimensions of multiword items between the experimental group and the control group?

RQ2: To what extent was the learning effectiveness through repeated oral reading influenced by students' vocabulary knowledge, dimensions of vocabulary, reading speed, and number of words per multiword unit?

Method

Participants

Sixty-two 18-19 year-old university freshmen majoring in hospitality and finance took part in the study. These students were randomly grouped by the university by their majors, but all of them were required to take Freshman English as one of their general education courses. Chinese was their native language. To measure students' sight and spoken vocabulary knowledge, 30 target words that are used in the first 1,000-word level of the New Vocabulary Levels Test (NVLT, Webb, Sasao, & Balance, 2017)[1] were selected to test the students' productive spoken forms and meanings. The students were allowed 30 seconds to orally read the selected words, and another 10 minutes to write the meanings of the words. Each correct item was awarded one point. The test results were used to classify students' vocabulary knowledge into three levels. Those who answered an average of 25 or more items correctly on both tests were classified into the high group, between 18 and 24 items the intermediate group, and below 18 the low group. Results showed that 25 students were in the high group (n =13, 12), 19 in the intermediate (n = 13, 6), and 18 in the low (n = 12, 6). The numbers in parentheses were the number of students in the experimental and control groups respectively.

Study Materials

Grouping idioms into thematic categories makes idioms easier to learn and memorize than unrelated ones (Boers, 2000; Cooper, 1998), so the target items for this study were all about the same theme. The reading text was selected from *Idiom Magic* by John Ryan (1994). *Idiom Magic* was written based on different themes using stories, and the theme—*time*—was selected for the present study (see Appendix A for the whole text). The text analyzed by BNC/COCA, shows that there are 332 words (8 words are contracted forms, so it was 324 words in total), with 122 word families and 158 word types (see Table 1). Despite the text being short, it contains many multiword units, most of which are related to "time," such as *once upon a time*, *on time*, *in time*, *the time of one's life*, *for decades*, and *for ages*. The text was divided into three paragraphs for students to do oral reading practice, and each paragraph contained 117, 96, and 111 words respectively (see Appendix A).

Table 1. Text Analysis

Word levels	Word counts		Word types		Word families
1,000	310	93.37%	144	91.14%	114
2,000	5	1.51%	4	2.53%	3
3,000	2	0.6%	2	1.27%	2
4,000	1	0.3%	1	0.63%	1
5,000	2	0.6%	2	1.27%	2
Proper nouns	0	0%	0	0%	0
Off list	12	3.61%	5	3.16%	
Total	332		158		122

---

[1] New Vocabulary Levels Test was updated from the previous versions of Vocabulary Levels Test (Nation, 1983, Schmitt, Schmitt, & Clapham, 2001), and added items for the first 1,000 level (the most frequent 1–1000 word families), which account for around 65–85% of spoken and written English.

Target Multiword Items

To select the target unknown MWIs, a pre-test containing 24 MWIs embedded in sentences (different sentences from those in the delayed post-test) was administered to all the student participants. The pre-test contained two steps. The first step was listening for form (spelling) and aural meanings. Students were asked to spell the idioms they heard and give a corresponding Chinese meaning for each item. The second step was to translate the target items (presented in isolation) into Chinese and make a sentence if they could. They also had to circle the unfamiliar words. After marking, 9 items were excluded because some items contained unfamiliar words, such as *for decades,* and some items were answered correctly by all or the majority of the students, e.g., *have a hard time*, *have a good time*, and *on time*. Finally, 15 target multiword expressions were selected based on the following criteria: They were items for which no students gave full correct answers for aural-form meaning and written-form meanings; every item was related to "time;" each unit contained no unfamiliar single words; every target item was listed in the Corpus of Contemporary American English (COCA); and all items appeared only once in the text. An analysis of all the words used in the 15 target items showed that all words were from the first 1,000 level. Among the 15 MWI, nine items were from the first paragraph, three from the second, and three from the third paragraph. The length (the number of words) of an MWI may also affect students' acquisition of it. Among the 15 selected MWIs, four items have 2 words, three items 3 words, five items 4 words, and three items 5 words. The complete profile of the target MWIs is presented in Table 2.

Table 2. Profile of the Target Multi-word Items

#	Target items	Target item at paragraph position	Words per item	Frequency in COCA
1	mark time	2	2	65
2	day in and day out	1	5	981
3	make up for lost time	3	5	214
4	keep time	1	2	167
5	from time to time	1	4	8,091
6	take one's time	1	3	4,984
7	run out of time	2	4	572
8	days are numbered	1	3	290
9	for ages	1	2	2,277
10	the time of one's life	3	5	129
11	golden years	2	2	521
12	around the clock	1	3	1,628
13	once in a while	1	4	6,647
14	make the big time	3	4	2,866
15	for the time being	1	4	9,082

Dependent Measures

Four dimensions of multiword knowledge were assessed. Like the pre-test, two forms of delayed post-test were developed: Form A measured students' knowledge in form and aural meaning recall. The sentences were pre-recorded by a native speaker of English. The students heard a sentence containing one of the target MWIs; the target item was then repeated once. Take one sentence for example: *Once upon a time, a rabbit and a tiger lived happily in the forest*. The students had to write down the repeated MWI (once upon a time) and the L2 meaning (從前). Form B tested written-form meaning and use. On the test sheet, students saw the 16 target items in isolation (one was a distractor), and they had to translate each item into Chinese. Finally, the students had to select one of the target items, changed its correct forms if necessary, and correctly put it in the blank of a sentence. An example is given. Students saw a target item *once in a while*. They translated it to Chinese meaning 偶而, then they had to correctly use the item to complete the sentences.

Marking and Scoring

A binary scoring system was adopted because of the analysis method (GLMM) that accepts only whole numbers; therefore, "1" was coded to the correct items and "0" to the incorrect items. Due to the binary scoring, strict marking (correct or wrong) was adopted. The marking was first done by two teaching assistants and a native English speaker, and then further checked by the researcher. If there was any inconsistency in the scores given, the researcher served as arbiter to decide the scores.

Spelling mistakes were not awarded points. For example, if "once upon a time" was spelled as "one's upon a time" or "for the time being" as "for the time been" or …bing," or "for ages" as "for age," then no point was given. Fifteen seconds was allowed for students to spell each of the MWIs and write down their meaning. All possible meanings were taught in the reading text before the test. For example, *once in a while* can be translated as 偶而 or 有時候. Either translation was awarded one point. In their written meaning recall, the meaning had to correspond to the context of the text.

Data Analysis

SPSS version 25 for Windows was used to analyze the data. The test was comprised of 15 items measured in four dimensions, adding up to 60 items in total. The internal reliability of the test was high (Cronbach's α = .95). The first research question compared the four dimensions of MWIs between the experimental group and the control group; one-way ANOVA was performed, and the effect sizes using Cohen's *d* were calculated. The second research question focused on the performance of the experimental group. The data were checked and entered in the long format; generalized linear mixed models (GLMMs) were used because the study involved an estimate of fixed and random effects and the dependent and independent variables involved repeated measures. The use of GLMMs also allows one to analyze all variables at one time, which can prevent Type I errors

The dependent variable was the score for each dimension of the target multiword units, and each score was binary (0 for an incorrect answer and 1 for a correct answer). The random factor was the student participants, and the fixed factors included the following: vocabulary knowledge (3 levels: high, intermediate, and low); oral reading rate (continuous variable);

dimensions of MWI knowledge tested (4 levels: spelling, aural meaning recall, written meaning recall, and use); and length of the MWI (2 levels: 2-3 words, 4-5 words).

Procedure

All participants were first given an oral reading and meaning translation test of a list of 30 words and then a pre-test on a list of 24 multiword expressions. From week 2 to week 4, students were asked to read a 324-word story divided into three paragraphs. To ensure the students would comprehend the content, each week, the students were explicitly taught one paragraph. The teaching focused mainly on explaining the meanings of the idioms and grammar (especially the tense). One important reason for adopting direct teaching was that many L2 learners do not recognize a string of known words as an idiom or individual words (Kim, 2016; Park & Chong, 2019). The participants of the present study also had this difficulty.

After formal instruction of a paragraph, the experimental group then orally read the paragraph five times on their own, and the time spent for each reading was recorded through students' cellular phones. After reading five times on their own, they read the paragraph to their teaching assistants. The time for the final reading was recorded by the teaching assistants and was used as one of the fixed variables (reading speed). The control group silently read another short story after each formal instruction. No teaching was given in week 5 due to a holiday. When the students returned to class, they were tested for spelling and aural meaning on the first session, then for written meaning and use on the second session. Students were not informed of the two-week delayed post-test in advance. The consent forms were signed by the students after the researcher explained the teaching and research purpose. See Table 3 for the summary of the procedure.

Table 3. Week-by-week procedure

Time	Tasks
Week 1	Students were given an oral reading and translations test on 30 1,000-word-level single words A pre-test was then administered for selecting the target multiword units.
Week 2-4	One paragraph was taught each week to all participants in the both groups, and then the students in the experimental group orally read the passage five times, then read the paragraph to the teaching assistants. The control group silently read a short story each time.
Week 5	Break
Week 6	A two-week delayed test was administered on four dimensions of multiword items in a sequence of spelling, oral meaning recall, written meaning recall, and use. Students signed the consent form after the researcher explained the purpose of the study.

Results

The descriptive statistics of the four dimensions of MWI in the two-week delayed post-test for the experimental group and the control group are presented in Table 4. As shown, the students in the experimental group scored an average of 72%, 51%, 28%, and 72% in the recall tests of spelling, aural meaning, use, and written meaning, respectively, whereas the scores were 50%, 8%, 18%, and 19% for the control group in the four dimensions of performance. With an exception in the dimension of use (p = .053), the differences between the experimental and control groups were statistically significant, and the effect sizes (Cohen's d) are large, 1.03, 1.90, and 2.47 for spelling, aural meaning, and written meaning (see Cohen, 2013; Plonsky & Oswald, 2014).

The results provide an answer to the first research question as to whether repeated oral reading six times improved students' retention of written meaning recall, aural meaning recall, and spelling to significant extents as compared to the control group. How to correctly use the multiword items was found to be the most difficult aspect even after orally reading six times.

Table 4. Descriptive Statistics (Maximum Score = 15) for the Experimental Group (EX) And the Control Group (CL) On the Tests of Spelling, Aural Meaning, Use, and Written Meaning

Dimensions	Subgroups	Mean (%)	SD	SE	95% CI Low – High	p	Effect size (Cohen's d)
Spelling	EX (n =38)	10.79 (72%)	3.07	.50	9.78 – 11.80	<.001	1.03
	CL (n = 24)	7.54 (50%)	3.23	.66	6.18 – 8.91		
Aural meaning	EX (n =38)	7.66 (51%)	3.84	.62	6.39 – 8.92	<.001	1.90
	CL (n = 24)	1.25 (8%)	2.85	.58	.05 – 2.45		
Use	EX (n =38)	4.26 (28%)	3.07	.50	3.25 – 5.27	n/s	0.52
	CL (n =24)	2.71 (18%)	2.94	.60	1.47 – 3.95		
Written meaning	EX (n =38)	10.87 (72%)	2.97	.48	9.89 – 11.84	<.001	2.47
	CL (n =24)	2.92 (19%)	3.46	.71	1.45 – 4.38		

To answer the second research question, GLMM was performed. Except for oral reading speed, the other three variables were found to have significant effects on retaining MWIs (see Table 5). For prior vocabulary knowledge, the odds ratio is statistically significant for the high level versus the low level (OR = 3.63, 95% CI = [1.77–7.54], p <.005), which indicates that the possibility of giving a correct response for students in the high level is 263% higher than those in the low level. No significant differences were found between other levels. In each prior vocabulary level, the retention rates on two-week delayed post-test were 64%, 57%, and 47% respectively (see Table 6).

Table 5. Summary of the GLMM

Predictor	B	OR (95%CI)	t	p
Intercept	1.02	2.76 (0.77 – 9.87)	1.57	.12
Prior vocabulary knowledge				
high vs low	1.29	3.63 (1.77 – 7.54)	3.52	.000 ***
intermediate vs low[a]	0.66	1.93 (0.94 – 3.95)	1.79	.074
high vs intermediate[a]	0.63	1.89 (0.91– 3.93)	1.69	.091
Dimensions				
spelling vs written meaning	2.26	0.97 (0.68 – 1.39)	-0.17	.867
aural meaning vs written meaning	2.23	0.33 (0.25 – 0.44)	-7.86	.000 ***
use vs written meaning[a]	1.16	0.11 (0.07 – 0.15)	-12.32	.000 ***
spelling vs use[a]	-1.16	9.55 (6.67 – 13.68)	12.32	.000 ***
aural meaning vs use[a]	1.10	9.27 (6.13 – 14.00)	10.57	.000 ***
spelling vs aural meaning[a]	1.07	3.18 (2.27 – 4.46)	6.71	.000 ***
Speed	-0.01	0.99 (0.98 – 1.00)	-1.56	.118
Number of words per phrase				
2-3 words vs 4-5 words	0.78	2.18 (1.78 – 2.66)	7.57	.000 ***

Note: [a] pairwise comparison, ***$p < .001$

In the four dimensions of multiword knowledge (Table 6), the results showed that if "written meaning" is used as the reference, the odds ratio for giving correct answers for written form meaning is significantly higher than for aural meaning (OR = 0.33, 95% CI = [0.25–0.44], $p < .005$), and use (OR = 0.11, 95% CI = [0.07–0.15], $p < .001$). When 'use' was the reference, students performed significantly better on spelling (OR = 9.55, 95% CI= [6.67–13.68], $p < .005$), and aural meaning (OR = 9.27, 95% CI = [6.13–14.00], $p < .005$). When we compared spelling with aural meaning, we found that students did significantly better on spelling than on aural meaning (OR = 3.18, 95% CI= [2.27–4.46], $p < .005$).

Table 6. The Mean Score for Each Dimension Among Students with High, Intermediate, and Low Vocabulary Knowledge in the Experimental Group

	Spelling (SD)	Aural meaning (SD)	Use (SD)	Written meaning (SD)	Total (SD)
High	.81 (.39)	.56 (.50)	.39 (.49)	.79 (.41)	.64 (.48)
Intermediate	.71 (.45)	.52 (.50)	.30 (.46)	.73 (.45)	.57 (.50)
Low	.63 (.49)	.44 (.50)	.16 (.36)	.64 (.50)	.47 (.50)

For the effect of oral reading speed on the retention of the MWIs, oral reading speed (B = -0.01, p = .118) did not significantly affect the retention of MWIs. Finally, the number of words in each MWI is found to have a significant effect on the retention of MWIs (OR = 2.18, 95% CI = [1.78–2.66], p < .005). The possibility of answering the items containing 2-3 words was 118% higher than the items with 4-5 words.

Discussion

The Effects of Repeated Oral Reading on Retention of Different Dimensions of MWIs

The results of the two-week delayed post-tests for the repeated oral reading group showed that students performed in the order of written meaning recall > aural meaning recall > aural form recall > use. Discussion for each dimension is as follows:

Production of aural forms

Orally repeating the text six times resulted in 22% points higher for the experimental group (72%) than for the control group (50%) in producing the correct forms of the MWIs. The difference between the two groups was not particularly large in aural form performance, but was very large between the present study and that of Alali and Schmitt's (2012). The reason could be that all MWIs in the present study were made up of known words, so that even students in the control group (who did not repeat the words as they learned them) were capable of spelling half of the MWIs correctly. In the study by Alali and Schmitt (2012), students retained only 0.5/10 items (5%). The reasons night be that their oral reading was done as a group, orally reading 10 times in 10 minutes, which was very demanding. More challenging still, each of their MWIs also contained unknown words. Most of the previous studies focused on written form recall rather than aural form recall; it is hence difficult to compare the differences. A general pattern, however, shows that productive form is in general more difficult than other dimensions, e.g., productive meaning recall (Alali & Schmitt, 2012; Chang & Chen, 2022; Puimege & Peters, 2019; Teng, 2019).

Although all MWIs in the present study were made up of familiar single words, students' test papers revealed that when these items were presented in spoken form, many mistakes were evident. For example, students may be more familiar with the singular form of *age* as in, *what is your age* than the plural form as in, *I have been waiting for you for ages*. The changes of form and meaning make it more difficult to process. Many students also substituted a word with a similar pronunciation. For example, *one's* was spelled for *once* in *once upon a time* or *once in a while* and *being* for *been* or *bing*, *lost* as *loss* or *losed*. These mistakes indicate the challenge for L2 learners to produce spoken forms and aural meanings at the same time.

Production of aural and written meaning

To the aural and written meaning recall, a larger effect size was found for written meaning. The experimental group outperformed the control group by 43% (51% vs 8%) points in producing the aural meanings of MWIs. It is interesting to note that the control group scored the lowest in aural meaning among the four dimensions. The score might imply that understanding the spoken meaning of an MWI was the most challenging aspect among the four dimensions. A few reasons may explain the phenomenon. Firstly, aural meaning test was given at the same time as aural form production. This meant that students' attention was divided between the two tasks. Secondly, unlike producing aural forms of MWIs, which were made up of known words, the meanings of the target MWIs were entirely unfamiliar to them before the treatment. Without repeatedly encountering the newly learned items, the retention rate was low. Finally, both tasks required students to *produce* the correct answers for the target items rather than selecting correct answers from a few options. For these above reasons, it may be safe to say that repeated oral reading six times improved aural meaning retention by 43% points higher compared to not having repeated oral reading. Among the limited studies that evaluated both written and aural meanings, students tended to score lower for aural meaning, and the retention rate corroborated that of Chang and Chen (2022).

Compared to the target items presented in written form, the experimental group outperformed the control group by 53% (72% minus 19%) points. Both groups scored higher for written meaning recall than for aural meaning recall. This may imply that the spoken form of multiword items is more difficult to comprehend than their written forms due to the spoken form being fleeting in nature whereas repeated oral reading can substantially improve learners' performance. The results corroborated Teng's (2019) findings on learning MWIs from captioning viewing, and Webb and Chang (2015) on reading while listening to short stories.

Among the fifteen target MWIs, the most confusing item was *golden years*. In Chinese, the best period of one's life literally translates to "golden age", so most of the participants intuitively translated the item into golden age. In addition, *for the time being* and *from time to time* were considered the most difficult items because many students left this item unanswered. Although the frequency in COCA for the two items were 9,082 and 8,091 respectively (see Table 2), the two items were difficult to remember. The reasons might be that the two items were abstract. Abstract words tend to be more difficult to learn than concrete ones (Laufer, 1990;). Remembering abstract items may require some strategies, and repeated oral reading may not be helpful to retain abstract items for long.

Use

It was not surprising that fewer than one-third of the target items were used correctly by the experimental group, and fewer than one-fifth by the control group. The difference between the two groups in how to use the multiword items is the smallest, compared to spelling, aural, and written meaning recall. One main reason could be that learning how to use a newly acquired item in a proper context takes time, and L2 learners may also require several encounters in different contexts, so they can become confident in using it. The result was similar to that of Chang and Chen (2022), in which a consistent pattern was shown that the knowledge of vocabulary in use scored the lowest in the one-week delayed post-test. The low correct use of these MWIs may imply that repeated oral reading six times helps L2 learners to comprehend meanings more than how to use the target multiword items. Although no results from previous

studies can be compared with the present one, many previous reading studies on frequency of encounter have demonstrated that different aspects of vocabulary knowledge might require a different number of encounters (Chang, 2019; Pellicer-Sânchez & Schmitt, 2010; Webb, 2007). The other reason was that the task also required students to produce correct grammatical use. For example, in a sentence— *your days were numbered*, if one wrote *your days was numbered* or *your day were numbered*, neither was awarded any points.

Taken altogether, comparing the above results to previous studies that examined learning of multiword units through repeated oral reading, the experimental group in the present study seemed to show much higher acquisition rates than those in Alali and Schmitt's (2012). A few factors may explain the results. Firstly, all single words in each item were known to the participants. The results seemed to support Zyzik (2011), whose students scored significantly higher on the items with known words than those items containing unknown words on the production test though not on the recognition test. Secondly, these multiword items had been explicitly taught and orally practiced six times before students took the two-week delayed post-test, so the acquisition rates seemed satisfactory. Thirdly, the present study required students to do individual repeated oral reading rather than choral reading, which ensured that every student did the practice on their own. Fourthly, collocations comprising node words with a higher number of collocates were found to be easier to learn than those with fewer collocates (Webb & Kagimoto, 2011). *Time* is a high-frequency word, and it has many collocates. Finally, all selected items are about *time*; thematically grouping target items together might have made the learning easier, or at least easier to guess the meaning in context (Zyzik, 2011).

Prior Vocabulary Knowledge Levels and the Retention of Different Dimensions of Vocabulary

The results of the present study corroborated those previous studies that took into account the role of prior vocabulary knowledge on learning single words (e.g., Webb & Chang, 2015; Zahar et al, 2001) and MWIs (Peters, 2016; Peters & Webb, 2018), in which a positive relationship between learners' prior linguistic knowledge and learning gains have been confirmed. In the present study, students' scores in the oral and written meaning of the 30 items selected from NVLT may further support those findings. The higher-level students were those who scored 25 or more out of 30 words selected from the most frequent 1,000-word level. Given this, all the individual words used in the target MWIs were likely to be familiar to them, so, they had a greater potential to comprehend the text better (Laufer & Ravenhorst-Kalovski, 2010), and to read the text more fluently (Chang, 2019), and could allocate their attention to the meanings of unknown MWIs (Peters & Webb, 2018). The lower-level students might have spent too much time decoding the meanings and the spoken forms of the individual words and giving less attention to the forms and meanings of the MWIs. Although the focus of the present study was on the MWIs, the results were consistent with those studies looking at single words (e.g., Peters & Webb, 2018; Webb & Chang, 2015).

To the gains of different dimensions, regardless of prior linguistic knowledge, it is interesting to note that across levels, students' retention rates for multiword items seemed to be quite consistent; all students scored similarly for spelling and written form meaning. The difference between the two dimensions was consistently within 2% across levels, which may suggest a close relationship between learning forms and written meanings (See Table 6). The mean score for aural meaning was consistently lower than for spelling and written meaning. The results

implied that the student participants had more difficulties decoding the aural meaning of a language string.

Oral Reading Speed and the Retention of Four Dimensions of Vocabulary

Analysis via GLMM showed that oral reading speed did not have a significant effect on the retention of multiword units. One of the main reasons could be that the effect of oral reading speed was moderated by the effect of prior vocabulary knowledge because the students possessing more vocabulary knowledge read more fluently than those who had lower-level vocabulary knowledge. As per de Jong and Perfetti (2011), fluency is like general proficiency. Higher-level students read faster and more accurately. If we look at the oral reading speed in each of the vocabulary knowledge levels, the reading speeds were 151, 124, and 105 wpm for the high-, the intermediate-, and the low-level students respectively (see Appendix B). There was a significant difference between each subgroup if we looked only at the variable of repeated oral reading without considering others. One-way ANOVA also confirmed that the more fluent readers read significantly faster than the lower-level students, $F (2, 149) = 124.56$, $p <. 001$. Three previous studies that adopted oral reading techniques in learning MWIs did not take into account oral reading speed as a factor that might affect learning rates (Alali & Schmitt, 2012; Chang & Chen, 2022; Durrant & Schmitt, 2010). The finding of this study is preliminary; more research into this area is needed.

Number of Words in Each Multiword Item and the Retention of Four Dimensions of Vocabulary

A significantly higher gain for fewer-word items was found in this study. The difference in the number of words for target items apparently influenced the learning to a significant extent. Looking at the effect of the number of words on acquiring MWIs more closely, we found that the acquisition rate for two-word multiword items was 70%, followed by three-word items (55%), and four-word items (52%), and five-word items (43%). The retention rate seemed to follow a very consistent pattern: the more words there were in an item, the lower the retention rate was, and the result corroborated the earlier studies that looked at the effect of single word length on learning outcome (Baddeley et al., 2002; Campoy, 2008). They generally found that shorter words were better recognized than longer ones, and the longer a word, the more room there was to err (Ellis & Beaton, 1993). Up to the present, most research focused on the properties of MWIs, such as frequency, predictability, and fixedness (Siyanova-Chanturia & Omidian, 2020). How the number of words in MWIs affects L2 acquisition and retention has rarely been explored; therefore, research into this area is warranted.

Pedagogical Implications and Conclusion

This study investigated the retention rates of high-frequency multiword items by 62 EFL students. The two-week retention rates seemed satisfactory for the experimental group. The results shed more light in that a large quantity of multiword items can be learned through explicit instruction (Laufer & Girsai, 2008; Webb & Kagimoto, 2009, 2011) and further enforced by repeated oral reading (Durrant & Schmitt, 2012). The results may have some important implications for teaching:

1. Doing repeated oral reading a few times can substantially improve students' recall of aural meaning and written meaning but not for correct use; therefore, additional exercises may be needed to enforce this dimension of knowledge.

2. It may be useful to ask students to orally read the text a few times after explicit instruction. By doing so, the opportunities for encountering the same items increase, and the retention rate may be improved.

3. Repeated oral reading might an effective way to draw learners' attention to forms (Gibson, 2008). When they encounter an item, such as *for ages* and *for decades*, in silently reading a text, the plural form of *age or decade* can be easily ignored. However, when students must orally read the item, they may notice the variation of pronunciation and form more easily.

4. Phrases with fewer constituent items were found to be easier to remember than more-word items. Given this, the lower-level students perhaps could start from learning phrases with fewer words first.

This study provides some empirical evidence that repeated oral reading can be used as a practice vehicle for retaining multiword items. Some limitations of the study should be pointed out before concluding the paper. Firstly, this study compared only a group doing repeated oral reading with a group that did not. Whether the effects of repeated oral reading could be superior to other types of practice technique is unclear. This area might be worth exploring in future studies. Secondly, this study did not contain an immediate post-test. The main reason was to reduce the learning effect from taking the same test again because some studies have recently shown higher scores in the delayed post-test (e.g., Peter & Webb, 2018; Webb & Chang, 2022). Some scholars suggest dividing the participants into two groups, one group taking the immediate post-test and the other group taking the delayed post-test (Webb & Chang, 2022). This solution however requires a larger number of participants and is a limitation of the present study. Thirdly, this study adopted a strict marking scheme, which limited our understanding of partial learning, especially for the lower-level students. Take one of the items for example – *One's days are numbered*. For many Chinese learners of EFL, understanding the meaning of this expression might be easier than spelling it because *day**s*** and *number**ed*** require both listening competence and syntactical knowledge to spell them correctly. How students process an item from entirely unknown to partially known and then fully known requires a more subtle marking scheme, and is worthwhile to explore in future studies. How to teach and assist L2 learners to efficiently acquire high-frequency multiword items will require more effort and creativity from researchers and language practitioners.

Finally, in addition to using repeated oral reading as a practice technique to enforce L2 learners' retention of MWIs, repeated oral reading can also be used for other purposes. For example, in L2 contexts, L2 learners generally lack opportunities to practice what they learn in real-world situations; repeated oral reading then can be used as a vehicle for them to practice their spoken skills. Repeated oral reading in fact "brings the language to life through voice" (Chang, 2019), which may add more variety to L2 learning and make L2 learning more enjoyable (Chang, 2019; Shimono, 2019).

About the Author

Anna C-S Chang is Professor in the Department of Applied English at Hsing Wu University, New Taipei, Taiwan, teaching courses on English listening, reading and vocabulary. Her main research interests focus on listening and reading development, and vocabulary learning. She has published extensively with internationally refereed journals. **ORCID ID**: 0000-0003-4216-6394

To Cite this Article

Chang, A. C-S. (2023). The effects of repeated oral reading practice on the retention of high-frequency multiword items for EFL learners: Multiple dimensions. *Teaching English as a Second Language Electronic Journal (TESL-EJ), 26* (4). https://doi.org/10.55593/ej.26104a9

References

Alali, F.A., & Schmitt, N. (2012). Teaching formulaic sequences: The same as or different from teaching single words? *TESOL Journal*, *3*(2), 153-180. https://doi.org/10.1002/tesj.13

Baddeley, A. (2013). *Essentials of Human Memory* (classic edition). Psychology Press. https://doi.org/10.4324/9780203587027

Baddeley, A. D., Chincotta, D., Stafford, L., & Turk, D. (2002). Is the word length effect in STM entirely attributable to output delay? Evidence from serial recognition. *Quarterly Journal of Experimental Psychology Section A, 55*(2), 353–369. https://doi.org/10.1080/02724980143000523

Boers, F. (2000). Metaphor awareness and vocabulary retention. *Applied Linguistics, 21*(4), 553–571. https://doi.org/10.1093/applin/21.4.553

Campoy, G. (2008). The effect of word length in short-term memory: Is rehearsal necessary? *The Quarterly Journal of Experimental Psychology, 61*(5), 724–734. https://doi.org10.1080/17470210701402364.

Chang, C- S (2019). The effects of repeated oral reading practice on EFL learners' oral reading fluency development. *The Reading Matrix, 19*(2), 103-113. https://readingmatrix.com/files/21-r9k89pn8.pdf

Chang, C-S. & Chen, T-Y. B. (2022). Comparing the multidimensional Effects of oral versus written exercises on the retention of L2 Vocabulary. *The Reading Matrix, 22*(2), 13-34. https://www.readingmatrix.com/files/29-cqhziqsn.pdf

Chang, C-S (2019). Effects of narrow reading and listening on L2 vocabulary learning: Multiple dimensions. *Studies in Second Language Acquisition, 41*(4), 769-794. https://doi.org/10.1017/S0272263119000032

Cohen, J. (2013). *Statistical Power Analysis for the Behavioral Sciences*. Routledge. https://doi.org/10.4324/9780203771587

Cooper, T. (1998). Teaching idioms. *Foreign Language Annals, 31*(2), 255-266. https://doi.org/10.1111/j.1944-9720.1998.tb00572.x

de Jong, N., & Perfetti, C. A. (2011). Fluency training in the ESL classroom: An experimental study of fluency development and proceduralization. *Language Learning, 61*(2), 533-568. https://doi.org/10.1111/j.1467-9922.2010.00620.x

Durrant, P., & Schmitt, N. (2010). Adult learners' retention of collocations from exposure. *Second Language Research, 26*(2), 163–188. https://doi.org/10.1177/0267658309349431

Dziemianko, A. (2010). Paper or electronic? The role of dictionary form in language reception, production and the retention of meaning and collocations. *International Journal of Lexicography, 23*(3), 257–273. https://doi.org/10.1093/ijl/ecp040

Ellis, N. C., & Beaton, A. (1993). Psycholinguistic determinants of foreign language vocabulary learning. *Language Learning, 43*(4), 559-617. https://doi.org/10.1111/j.1467-1770.1993.tb00627.x

Erman, B., & Warren, B. (2000). The idiom principle and the open choice principle. *Text, 20*(1), 29–62. https://doi.org/10.1515/text.1.2000.20.1.29

Fuchs, L., Fuchs, D., & Hosp, M. (2001). Oral reading fluency as an indicator of reading competence: A theoretical, empirical and historical analysis. *Scientific Studies of Reading, 5*(3), 239-256. https://doi.org/10.1207/S1532799XSSR0503_3

Gathercole, S. E., & Conway, M. A. (1988). Exploring long-term modality effects: Vocalization leads to best retention. *Memory & Cognition, 16*(2), 110-119. https://doi.org/10.3758/BF03213478

Gibson, S. (2008). Reading aloud: a useful learning tool? *ELT journal, 62*(1), 29-36. https://doi.org/10.1093/elt/ccm075

Grabe, W. (2009). *Reading in a Second Language: Moving from Theory to Practice.* Cambridge University Press. https://doi.org/10.1017/CBO9781139150484.013

Gyllstad, H., & Wolter, B. (2016). Collocational processing in light of the phraseological continuum model: Does semantic transparency matter? *Language Learning, 66*(2), 296–323. https://doi.org/10.1111/lang.12143

Horst, M., Cobb, T., & Meara, P. (1998). Beyond A Clockwork Orange: Acquiring second language vocabulary through reading. *Reading in a Foreign Language, 11*, 207-223. https://eric.ed.gov/?id=EJ577617

Jiang, X.-Y. (2016). The role of oral reading fluency in ESL reading comprehension among learners of different first language backgrounds. *The Reading Matrix, 16*(2), 227-242. https://eric.ed.gov/?id=EJ1115275

Kim, C. (2016). L2 learners' recognition of unfamiliar idioms composed of familiar words. *Language Awareness, 25*(1-2), 89-109. https://doi.org/10.1080/09658416.2015.1122025

Kuhn, M., & Stahl, S. (2003). Fluency: A review of developmental and remedial practices. *Journal of Educational Research, 95*(1), 3-21. https://doi.org/10.1037/0022-0663.95.1.3

Laufer, B. (1990). What's in a word that makes it hard or easy: Some intralexical factors that affect the learning of words. *International Review of Applied Linguistics in Language Teaching, 28*(4), 293-308. https://doi.org/10.1515/iral.1990.28.4.293

Laufer, B. (2013). Lexical thresholds for reading comprehension: What they are and how they can be used for teaching purposes. *TESOL Quarterly, 47*(4), 867-872. https://doi.org/10.1002/tesq.140

Laufer, B. (2011). The contribution of dictionary use to the production and retention of collocations in a second language. *International Journal of Lexicography, 24*(1), 29-49. https://doi.org/10.1093/ijl/ecq039

Laufer, B., & Girsai, N. (2008). Form-focused instruction in second language vocabulary learning: A case for contrastive analysis and translation. *Applied linguistics, 29*(4), 694-716. https://doi.org/10.1093/applin/amn018

Laufer, B., & Ravenhorst-Kalovski, G. (2010). Lexical threshold revisited: Lexical text coverage, learners' vocabulary size and reading comprehension. *Reading in a Foreign Language, 22*(1),15–30. https://files.eric.ed.gov/fulltext/EJ887873.pdf

Lin, Y-T. (2016). Oral reading fluency with peer-assisted reading via telecollaboration. *Innovation in Language Learning and Teaching, 10*(3), 272-281. https://doi.org/10.1080/17501229.2015.1127924

Lindstromberg, S., & Boers, F. (2008). Phonemic repetition and the learning of lexical chunks: The mnemonic power of assonance. *System, 36*(3), 423–436. https://doi.org/10.1016/j.system.2008.01.002

Macis, M., & Schmitt, N. (2017). The figurative and polysemous nature of collocations and their place in ELT. *ELT Journal, 71*(1), 50–59. https://doi.org/10.1093/elt/ccw044

Nation, I.S. P. (1983). Testing and teaching vocabulary. *Guidelines, 5*, 12-25.

Nation, I. S. P., & Webb, S. (2011). *Researching and Analyzing Vocabulary*. Heinle.

Nguyen, T. M. H., & NWebb, S. (2017). Examining second language receptive knowledge of collocation and factors that affect learning. *Language Teaching Research, 21*(3), 298-320. https://doi.org/10.1177/1362168816639619

Park, J., & Chon, Y. V. (2019). EFL learners' knowledge of high-frequency words in the comprehension of idioms: A Boost or a burden? *RELC Journal, 50*(2), 219-234. https://doi.org/10.1177/0033688217748024

Pellicer-Sánchez, A., & Schmitt, N. (2010). Incidental vocabulary acquisition from an authentic novel: Do things fall apart? *Reading in a Foreign Language, 22*(1), 31-55. https://eric.ed.gov/?id=EJ887875

Peters, E. (2016). The learning burden of collocations: The role of interlexical and intralexical factors. *Language Teaching Research, 20*(1), 113-138. https://doi.org/10.1177/1362168814568131

Peters, E., & Webb, S. (2018). Incidental vocabulary acquisition through viewing L2 television and factors that affect learning. *Studies in Second Language Acquisition, 40*(3), 551-577. https://doi.org/10.1017/S0272263117000407

Pigada, M., & Schmitt, N. (2006). Vocabulary acquisition from extensive reading: A case study. *Reading in a Foreign Language, 18*(1), 1-28. https://files.eric.ed.gov/fulltext/EJ759833.pdf

Plonsky, L., & Oswald. F. L. (2014). How big is "Big"? Interpreting effect sizes in L2 research. *Language Learning, 64*(4), 878–912. https://doi.org/10.1111/lang.12079

Puimège, E., & Peters, E. (2019). Learning L2 vocabulary from audiovisual input: an exploratory study into incidental learning of single words and formulaic sequences. *The Language Learning Journal, 47*(4), 424-438. https://doi.org/10.1080/09571736.2019.1638630

Rasinski, T., Homan, S., & Biggs, M. (2009). Teaching reading fluency to struggling readers: Method, materials, and evidence. *Reading & Writing Quarterly, 25*(2-3), 192-204. https://doi.org/10.1080/10573560802683622

Ryan, J. (1994). *Idiom Magic*. Overseas Radio & Television Inc. Taipei, Taiwan.

Shimono, T.R. (2019). The effects of repeated oral reading and timed reading on l2 oral reading fluency. *The Reading Matrix, 19*(1), 139-154. https://readingmatrix.com/files/20-0207e343.pdf

Siyanova-Chanturia, A., & Omidian, T. (2020). Key issues in researching multiword items. In S. A. Webb (Ed.) *The Handbook of Vocabulary Studies* (pp.529-544). Routledge. https://doi.org/10.4324/9780429291586-33

Teng, F. (2016). Incidental Vocabulary Acquisition from Reading and Listening: The Effects of Word Exposure Frequency. In E. Dorman, and J. Bidal (Eds.) *Departing from Tradition: Innovations in English Language Teaching and Learning* (pp. 82–207). Cambridge Scholars. https://doi.org/10.1080/17501229.2016.1203328

Teng, F. (2019). The effects of video caption types and advance organizers on incidental L2 collocation learning. *Computers and Education, 142*. https://doi.org/10.1016/j.compedu.2019.103655

Uchihara, T., Webb, S., & Yanagisawa, A. (2019). The effects of repetition on incidental vocabulary learning: A meta-analysis of correlational studies. *Language Learning, 69*(3), 559-599. https://doi.org/10.1111/lang.12343

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. https://doi.org/10.1017/S0272263105050023

Webb, S. (2007). The effects of repetition on vocabulary knowledge. *Applied linguistics, 28*(1), 46-65. https://doi.org/10.1093/applin/aml048

Webb, S., & Chang, C-S (2022). How does mode of input affect the incidental learning of collocations? *Studies in Second Language Acquisition, 44*(1), 35-56. https://doi.org/10.1017/S0272263120000297

Webb, S., & Chang, C-S (2015). How does prior word knowledge affect vocabulary learning progress in an extensive reading program? *Studies in Second Language Acquisition, 37*(4), 651- 675. https://doi:10.1017/S0272263114000606

Webb, S., & Kagimoto, E. (2009). The effects of vocabulary learning on collocation and meaning. *TESOL Quarterly, 43*(1), 55–77. https://doi.org/10.1002/j.1545-7249.2009.tb00227.x

Webb, S., & Kagimoto, E. (2011). Learning collocations: Do the number of collocates, position of the node word, and synonymy affect learning? *Applied Linguistics, 32*(3), 259–276. https://doi.org/10.1093/applin/amq051

Webb, S., Sasao, Y., & Ballance, O. (2017). The updated Vocabulary Levels Test: Developing and validating two new forms of the VLT. *ITL-International Journal of Applied Linguistics, 168*(1), 33-69. https://doi.org/10.1075/itl.168.1.02web

Zahar, R., Cobb, T., & Spada, N. (2001). Acquiring vocabulary through reading: Effects of frequency and contextual richness. *Canadian Modern Language Review, 57*(4), 541-572. https://doi.org/10.3138/cmlr.57.4.541

Zyzik, E. (2011). Second language idiom learning: The effects of lexical knowledge and pedagogical sequencing. *Language Teaching Research, 15*(4), 413-433. https://doi.org/10.1177/1362168811412025

Appendix A: The text

Paragraph 1 (117 words)

Once upon a time there was an old clockmaker named Hans. He made clocks day in and day out. His timepieces kept time very well. Usually, he wasn't too busy, so he could take his time. Once in a while, though, he had to make many clocks at one time. To finish them on time, he would work around the clock. But from time to time, he took time out to talk to his wife, Greta.

 "Greta," he said, "I've been making clocks year in, year out for ages. I'm getting tired of it. I'll keep working for the time being, but my clockmaking days are numbered. Soon the time will be ripe for me to retire."

Paragraph 2 (96 words)

His wife answered, "I agree, Hans. You've been making clocks day and night. It's about time you stopped working. I think you should retire in time to enjoy your golden years. You've had a hard time for decades. Now you deserve to have a good time before you run out of time."

So that's what Hans did. He told his customers one month ahead of time that he was going to close his store. During that last month, he didn't just mark time. He worked very hard to finish his last clock. The days flew by.

Paragraph (111 words)

After the last customer picked up the last clock, Hans looked around his store one last time. For the first time in 50 years, there was no sound of ticking clocks. Hans felt happy and sad at the same time. "I never made the big time," he thought to himself, "but my time was well spent." With that, he locked the door for the last time and went home. Hans and Greta took a long vacation. They wanted to make up for lost time. They traveled around the world.

They visited all of their relatives and friends. They had a great time. In fact, they had the time of their lives.

Appendix B: Oral Reading Speed (in Words per Minute) for Students with Different Levels of Vocabulary Knowledge

VK	Mean	N	SD	Minimum	Maximum
High	151	52	18	131	200
Intermediate	124	52	11	106	145
Low	105	48	14	77	126
Total	127	152	24	77	200