THE EFFECT OF IMAGERY ON MEANING AND FORM RECOGNITION AND RECALL OF IDIOMS

By

Marisol Masso

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Teaching English to Speakers of Other Languages—Master of Arts

2019

ABSTRACT

THE EFFECT OF IMAGERY ON MEANING AND FORM RECOGNITION AND RECALL OF IDIOMS

By

Marisol Masso

This research project aims at investigating whether the use of imagery – pictures and etymological notes – has a facilitative effect on the recognition and recall of the form and meaning of American idioms. Spanish speakers (n = 71) enrolled in English programs in Argentina participated in a pretest, a learning phase activity, and an immediate post-test. One week later, participants (n = 39) took a delayed post-test in-person or online. All tests had the same format: a sentence completion task, a receptive form multiple-choice task, a definition task, and a receptive meaning multiple-choice task. Results indicate that all conditions – including pictures, etymological notes, and the combination of both – had no statistically significant differences between them, and thus, they all contributed to idiom learning. However, the data also shows a tendency for pictures to facilitate structural elaboration of idioms in the short-term while etymological notes seem to hinder it.

ACKNOWLEDGEMENTS

I would like to give a special thanks to Prof. Godfroid who helped me draft this research project and encouraged me to materialize it. I would also like to thank the College of Arts and Letters that granted me a summer research fellowship to collect the data for this research study. I would like to show gratitude to two colleagues of mine, Andrés and Facundo, who collaborated in the last rounds of data collection in Argentina. I want to give special thanks to my advisor Prof. Polio for her support throughout the data collection process and her guidance in the thesis writing. I am also thankful for Prof. Spinner who provided useful feedback. Much appreciation also goes to Juna for her assistance in the statistical aspects of my study and for helping me rediscover the value of statistics. Finally, I would like to thank my family and friends for the words of encouragement and their unconditional support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1 Literature Review

Introduction

Formulaic language is a significant component of vocabulary. The importance of formulaic language in both written and spoken English lies in its abundance in the language and the different communicative purposes it realizes in language use (Nation, 2013; Martinez & Schmitt, 2012; Schmitt, 2010). Wray (2008) defines formulaic language as a morpheme equivalent unit (MEU). In this view, MEU consists of a word or word strings that are processed like a morpheme, rather than individual words, which are believed to make up an atomic lexicon. The formulaic expressions this study is interested in are, in Wray's (2008) words, true idioms, fully-fixed multiword strings that express meaning metaphorically. As Martinez and Schmitt (2012) noted, formulaic expressions vary in terms of their degree of compositionality. They explain that formulaic sequences with high compositionality, as in *at all times,* consist of individual words that match the meaning of the phrase, which in turn, can be easily understood from the component words. In contrast, formulaic language with lower compositionality- which is the focus of this study- is more figurative and less transparent in meaning. This is the case of expressions such as *at all*, whose meaning cannot be decoded by the literal reading of the component words, placing them in the right of the spectrum of idiomaticity proposed by Lewis (1993).

Formulaic language is characteristic of effective and fluent communication. It allows speakers to express and interpret a great deal of meaning efficiently with little processing (Martinez & Schmitt, 2012; Wray, 2008). When both native and non-native speakers of English make automatic use of accurate figurative language, they can free up their minds to perform other language processes (Martinez & Schmitt, 2012; Schmitt, 2010), and thus, are perceived as

fluent and proficient users of the language (Boers & Lindstromberg, 2009; Martinez and Schmitt, 2012). Additionally, Martinez & Schmitt (2012) claimed that the fluent use of idioms shows that the speaker shares cultural knowledge, which in turn, might contribute to higher social acceptance and integration of second language (L2) learners into the English-speaking community.

Adult L2 learners' efforts at learning formulaic language is, however, hampered by several difficulties. According to Wray (2008), one difficulty is the L2 learners' tendency to analyze formulaic sequences into individual lexical units. Boers & Lindstromberg (2009) noted that this inability to decode the meaning of the sequences as whole chunks reduces the reading or listening comprehension of texts since the speaker or listener is not able to interpret the idioms in context and overlooks the evaluative function they serve for the understanding of written or aural texts. The authors further explained that all idioms express evaluations – intrinsically linked to the culture of the speaking communities – that learners are not able to fully identify or express unless they have a wide knowledge of idioms. Another problematic tendency is that learners prefer to express meanings using an alternative non-native-like language, rather than formulaic sequences they have been exposed to (Boers & Lindstromberg, 2009; Wray, 2008). Adult L2 learners who avoid using formulaic language because they either lack the linguistic knowledge of the idiomatic expression or deliberately avoid using it – seem to use instances of paraphrase, non-idiomatic synonyms and other non-idiomatic phrases with alternate literal meanings (Irujo, 1993). One more difficulty is the arbitrary nature of the form-meaning connection of formulaic language with a lower degree of compositionality that hinders L2 learners' interpretation of these expressions (Martinez & Schmitt, 2012). Finally, some formulaic language also appears to be less accessible due to the abstract themes they map onto (e.g. *be beyond someone*) and less

memorable for lacking mnemonic potential i.e., some idioms are difficult to remember for their spelling, pronunciation, grammatical category, meaning, word associations, (Hulstijn, 2001, as cited in Boers & Lindstromberg, 2009) and the visual and motoric images their meanings elicit (Boers & Lindstromberg, 2009).

Teaching Formulaic Language

Pavičić Takač (2008) noted that pictorial elucidation is one of the most common ways of introducing new lexical items. Pictorial elucidation, also sometimes referred to as picture elicitation, consists of the use of pictures to foster receptive or productive word knowledge (Boers, 2013). Visuals can directly connect the meaning of a word to a concrete object or a phenomenon (Pavičić Takač, 2008). Gallese and Lakoff (2005) claimed that people use their imagination, i.e. mental simulation to learn a concept. They further explained that people understand the concept of a word such as *grasp* because they imagine the action of *grasping* (i.e. they can imagine themselves or others performing the action). The mental ability to visually imagine concepts uses some of the same areas of the brain that are used for seeing (Gallese & Lakoff, 2005). Thus, pictures serve to stimulate mental simulation and strengthen the link between what learners see and the mental picture of the concept.

However, learners usually forget the meaning of new words, and therefore, need to review new lexical items to store them in the long-term memory. Semantic elaboration is one of the vocabulary teaching strategies that can encourage learners to retrieve the meaning of words. Semantic elaboration, also referred as meaning-focused elaboration, is a process that contributes to the formation of a mental picture of the meaning of a word, creating stable memory traces (Boers & Lindstromberg, 2009). One semantic elaboration technique is dual coding (Paivio,

1986). Dual coding is a theory that explains how verbal and non-verbal (images, sounds, actions, sensations, emotions) mental systems process imagery and linguistic information. It can create semantic networks and stimulate deep processing, which can lead to memorization of the concept (Pavičić Takač, 2008). Dual coding can be stimulated through imagery, mime and physical enactment, and pictorial elucidation. This theory suggests that visual illustrations can help to facilitate comprehension and retention of new concepts -particularly concrete concepts- through the activation of knowledge and the stimulation of mental images in learners (Clark & Paivio, 1991). Idioms and figurative language have also been studied under the principles of dual coding, revealing that verbal and imaginal processes have contributed to the retention of the meaning of those expressions (Clark & Paivio, 1991; Sadoski & Paivio, 2013).

In Boers et al.'s (2009) study, pictorial elucidation stimulated dual coding by presenting verbal explanations of the origin of idioms (e.g. *the carrot and the stick* can be traced back to the scene of a farmer trying to persuade a mule by giving it either a carrot as a reward or a stick as a punishment) accompanied by pictorials representing the literal meaning of the individual words altogether. Pictures and etymological information seem to favor the formation of a mental picture associated with the literal meaning of the idioms (Lakoff, 1987). In fact, the processing of mental imagery – verbal and visual information – can result in deeper processing (Craik & Lockhart, 1972), and therefore, facilitate the recognition and recall of vocabulary. Boers et al. (2009) noted that special caution must be taken in the selection of the pictures since they must elucidate the meaning of the idioms by clarifying the literal origin from which the idiomatic meanings are derived (e.g. a picture of the carrot and the stick showing a mule that is promised a carrot if it moves forward, and simultaneously, threatened to be struck with a stick if it does not). Boers and Lindstromberg (2009) proposed that teachers should make targeted formulaic language,

specifically idioms, more memorable by engaging learners in semantic elaboration as well as structural elaboration. Structural elaboration is form-focused elaboration, meaning that it favors the memorization of the form of idioms by using phonological repetition such as alliteration (e.g. *it takes two to tango*).

Previous Empirical Research

The current study is grounded in previous experimental studies conducted to investigate the effect of verbal input and pictorial elucidation in the recollection of meaning and form of idioms (Boers, Lindstromberg, Littlemore, Stengers, & Eyckmans, 2008; Boers et al., 2009). Boers et al.'s (2008) experimental study suggested that verbal explanations facilitate the retention of the actual idioms themselves while pictorial elucidation seems to be beneficial for meaning retention. The use of pictures appears to foster receptive word knowledge, yet there is no substantial evidence to prove the same result applies to productive knowledge. Pictures have the potential to elucidate concepts, not words, which in turn, provides no certainty about the learner's ability to recall the precise form of the L2 idioms (Boers, 2013).

Boers et al. (2009) conducted a study based on Boers et al.'s (2008) study to examine the potential of pictures in form retention of idioms. Participants first solved a meaning multiple-choice test. This time an origin multiple-choice test with information about the origin of English idioms was delivered before the last test, a gap-fill exercise. Half of the participants were presented with only a verbal explanation of the origin of idioms while the other half was exposed to verbal explanations accompanied with pictures. The interesting result of this study is that the presence of pictures in the input had a negative impact on the learners' performance. Learners were able to remember the concepts elicited by the pictures, but they could not remember the

precise lexical composition of the idioms (Boers, 2013). In phrases like *keep a tight rein on someone* and *play second fiddle*, the word *rein* was replaced by *rope*, and *fiddle* by *violin*, to name a few. According to Nelson et al. (1976), this phenomenon is known as picture superiority in which pictures are better recalled than words when they are presented altogether (as cited in Boers, 2013). Boers et al. (2009) argued that since learners' attention appears to be drawn to the pictures, rather than the words, pictures seem to have a distracting effect on learners, resulting in the learners' inability to recall the precise linguistic form of L2 idioms.

Boers et al. (2009) acknowledges some weaknesses in the experimental design of this study that prevent generalization of the findings. First, only thirty-eight subjects participated in the study, which is a relatively low number of participants. More participants are necessary to balance out variations due to individual differences of the subjects (Schmitt, 2010). Second, immediate post-tests cannot inform us about the dynamic nature of vocabulary learning. Given that vocabulary learning is incremental in nature and subject to vocabulary attrition, immediate posttests should be employed to provide additional exposure, as well as delayed posttests to enhance learning, and thus, indicate long-term retention (Schmitt, 2010). Third, the use of unbalanced test formats to measure receptive and productive knowledge needs to be tackled in future research studies. This was the case in Boers et al.'s (2009) study in which receptive and productive knowledge were compared after being measured by different test formats; receptive knowledge was measured by recognition tests (a meaning multiple-choice test and an origin multiple-choice test) and productive knowledge was measured by one recall test (a gap-fill test). When comparing receptive and productive knowledge, a study should use equivalent test formats, which means including either recognition or recall items in both receptive or productive

tests (Nation, 2013). In sum, the gaps of this study reveal that the potential of pictorial elucidation as a tool to enhance the learning of idioms remains to be specified.

Further research on the use of pictorial elucidation was conducted by Szczepaniak and Lew (2011). Rather than analyzing pictures in vocabulary learning exercises as in previous research (Boers et al., 2008; Boers et al., 2009), the purpose of their study was to test the effect of imagery – pictorial illustrations and etymological notes – employed in idioms entries in idioms dictionaries. The picture selection was done on the basis that they were to illustrate the literal meaning of the idioms or one of their component words. The expression *sitting duck,* for instance, was illustrated by a picture of a duck that is sitting. An etymological note explained the origin of the same phrase as a duck being an easy target for hunters when it is sitting on the water or on the ground. A group of sixty-two Polish participants were randomly exposed to one of four entry types which provided different types of information (type 1: definition + example; type 2: definition + example + etymological note; type 3: definition + example + picture; type 4: definition + example + picture + etymological note). An immediate and delayed post-test were used to measure the immediate and delayed retention of the form and meaning of idioms. A recognition meaning multiple-choice test measured receptive knowledge, and a recall test with a given lexical component to elicit the full idiomatic form measured productive knowledge. This study reported similar results to Boers et al. (2008) and Boers et al. (2009) in that the meaning of the idioms was better recalled than the phrase itself. However, Szczepaniak and Lew (2011) revealed that only pictorial elucidation had a statistically significant effect on the form retention of idioms in both the short- and long-term. Similarly, only pictorial elucidation had a significant effect on meaning retention of idioms in the short-term. Yet, in the long-term, there was just a tendency to recall the meaning of idioms. These research findings merit further investigation to

determine the effect of pictorial elucidation and etymological explanations on the immediate and delayed recall and retention of the meaning and form of idioms.

Study (Rationale)

The aim of the present study is to shed light on the conflicting findings in Boers et al. (2009) and Szczepaniak and Lew (2011). Boers et al. (2009) reported striking findings about the negative effect of pictorial elucidation on form learning while Szczepaniak and Lew (2011) provided evidence for the beneficial effect of pictorial elucidation on the immediate and delayed retention of forms. The current study, therefore, is interested in answering the question of whether imagery – pictorial elucidation and etymological explanations – has a facilitative or inhibitory effect on the retention and recall of the meaning and form of idioms in the short and long-term. The purpose of this study is not a replication of the aforementioned studies. Rather, the current study is an extension of Szczepaniak and Lew (2011) in that it will strive to give a more complete picture of idiom learning. This study is focused on establishing the form-meaning connection of idioms and measuring the form-meaning recognition and recall of the targeted idioms (Nation, 2013). Therefore, the research uses receptive multiple-choice tests that measure the form-meaning link at a recognition level, and productive tests that measure form recognition and recall. First, the present study will investigate the role of imagery in the immediate recognition and recall of both meaning and form of idioms. Second, the study will also explore to what extent imagery can foster retention of meaning and form in the long-term. The value of this study lies in the analysis of how effective the use of imagery can be to recognize and recall the meaning and form of English idioms. Such findings could provide English as a second language (ESL) and English as a foreign language (EFL) teachers with valuable pedagogical implications

for teaching idioms, by gaining a better understanding of what teaching strategies can foster effective learning of idioms.

Research Questions

The research questions of this study are the following:

RQ1: Does pictorial elucidation enhance form and meaning recognition and recall of idioms in the immediate posttest?

RQ2: Does pictorial elucidation result in more durable retention of form and meaning recognition and recall of idioms in the delayed posttest?

RQ3: Do etymological notes enhance form and meaning recognition and recall of idioms in the immediate posttest?

RQ4: Do etymological notes result in more durable retention of form and meaning recognition and recall of idioms in the delayed posttest?

Chapter 2 Method

Participants

The participants include 71 Spanish native speakers from Argentina, ranging from an upper--intermediate to low--advanced level of English. They were enrolled in four-year English programs such as a BA in English Teaching, Translation, and English Philology at three universities and two colleges in the Midwest of Argentina. Table 1 shows background information of participants' gender, age, and first language.

Table 1

Background information

	Gender		Age group					L1
	F	M	18-19	20-23	24-26	27-29	+ 30	Spanish
N	60	11	21	24	13	9	4	71
%	85	15	30	34	18	13	5	100

Note. L1 = first language.

Materials

As mentioned, this study is an extension of Szczepaniak and Lew (2011), and as such, it uses some of the materials and instruments from their study. The targeted vocabulary included 24 American idioms, listed in Appendix A. Fifteen targeted items were used in Szczepaniak and Lew's (2011) study: *sitting duck; loose cannon; white elephant; sour grapes; grass roots; red herring; lame duck; red tape; ax to grind; the gloves are off; let your hair down; lock horns with someone; walk the plank; get cold feet; move the goalposts.* The other nine idioms along with examples and explanations of their origin were taken from the *US Department of State* official

website for English teachers for the design of the tests to expose students only to frequent authentic American idioms. They include: *bite the bullet; sacred cow; green thumb; last straw; acid test; pull strings; dark horse; jump the gun; pink slip.* More sentence examples were taken from the *Online Oxford Learner's Dictionary* and the *Online Cambridge Dictionary* to design the tests. The pictures of the idioms mentioned above were taken from the Internet via the Google image search function. All targeted items were selected for their potential difficulty in terms of opacity, ability to be illustrated, and lack of Spanish equivalents (Szczepaniak and Lew, 2011). In other words, I, a Spanish speaker, considered the chosen target vocabulary to be hard to guess by Spanish speakers (e.g. *loose cannon, red herring, lame duck*); the selected pictures illustrated the literal meaning of the idioms, reducing the risk of personal interpretation (e.g. a picture of *bite the bullet* depicts someone biting a bullet); and the targeted vocabulary lacked exact Spanish translations to prevent first language interference (e.g. the idiom *kill two birds with one stone* has an almost exact Spanish equivalent: *kill two birds with one shot*). The criterion for picture selection was quality and precision of the pictures to clarify the literal reading of the idioms, and thus, stimulate pictorial elucidation.

This study consisted of four phases: an untimed pre-test, exposure to the targeted idioms (10 minutes), an immediate post-test (20 minutes) and a delayed post-test (20 minutes). The same receptive multiple-choice tasks and productive recall exercises (sentence completion task and definition task) were used in each of the phases. In total, there were four different types of tasks, which were a) a sentence completion task, b) a receptive form multiple-choice task, c) a definition task, and d) a receptive meaning multiple-choice task. The first two tasks tested participants' productive and receptive knowledge of the form respectively, and were followed by the other two tasks that measured participants' productive and receptive knowledge of the

meaning of idioms accordingly. The tasks were placed in such a sequence to prevent a learning effect from the tasks, and thus, biased results. If the tasks measuring receptive knowledge had been placed first, they would have revealed the actual form and meaning of the idioms before having measured participants' productive knowledge of the form and meaning of the targeted idioms. An example of each task is given in (1) and all the tests can be found in Appendix B.

(1) a) Sentence completion task

The salesman has been trying to get rid of that car for more than a year. It costs too much to run and insure, so no one wants it—it's a ***w***_____ _____.

b) Receptive form multiple-choice task:

The theatre is a real_____. It costs millions to build and nobody ever goes there.

a) White lie
b) Wet blanket
c) White elephant
d) Wet whistle
e) I don't know

c) Definition task:

What is the meaning of the expression *"white elephant"*?

_____

d) Receptive meaning multiple-choice task:

1. White elephant

a) something very rare and precious that everyone wants to have
b) something that is useless and expensive to maintain
c) something imaginary, fictitious, hard to believe
d) someone who is socially awkward and does not have many friends
e) I don't know

Qualtrics was the online tool that displayed the exercises above. In the learning phase, the online website "Quizlet" and the activity "Flashcards" from this site was used.

To ensure the tasks used were valid and reliable, average item facility, average item discrimination, and Cronbach's alpha were calculated as shown in Table 2. Results indicate an appropriate item facility and item discrimination (no higher than 0.51) in all productive tasks. Items in the receptive form and meaning multiple-choice tasks at immediate time point showed high average item facility (0.94 and 0.90 respectively), and low item discrimination (0.13 and 0.20 accordingly). In other words, results for average item facility showed that items in the receptive tests were easier than in the productive tests. Results for item discrimination also showed that items in productive tests better discriminated word knowledge than receptive tests did. However, average item discrimination showed no negative values for receptive tests, which suggests that items in receptive tests reflect the constructs that each task was intended to measure. These figures are not completely surprising, because as Nation (2013) noted, it is easier to gain word knowledge receptively than productively. In other words, receptive tests are easier than productive tests. Therefore, the author explains, when conducting experimental comparisons of receptive and productive knowledge, both recognition and recall items should be included to make the study design balanced. For a study design to be balanced, it should measure both the receptive and productive knowledge of idioms as the present study did.

Cronbach's alpha was used to measure item reliability across tasks and time points. This reliability estimate displayed a high level of internal consistency among the individual test items in all tasks. Again, results were slightly lower for the receptive form multiple-choice test at the immediate time point, and the receptive meaning multiple-choice tests at the delayed time point,

yet they were still highly reliable (α = .816 and α = .897 respectively). Overall, all items reliably measured the constructs they were designed to measure.

Table 2

Reliability estimates of the tests

Test type	Task type	Time	Average IF (min-max)	Average ID (min-max)	Cronbach's α
Productive	sentence completion	IMP	0.47 (0.75-0.15)	0.41 (0.62-0.17)	.912
		DP	0.46 (0.77-0.21)	0.51 (0.70-0.36)	.947
	definition task	IMP	0.51 (0.94-0.27)	0.41 (0.63-0.11)	.917
		DP	0.47 (0.94-0.24)	0.43 (0.74-0.14)	.928
Receptive	FMC	IMP	0.94 (1.00-0.85)	0.13 (0.33-0)	.816
		DP	0.75 (0.97-0.41)	0.32 (0.58-0.05)	.911
	MMC	IMP	0.90 (0.97-0.77)	0.20 (0.47-0.05)	.900
		DP	0.74 (0.97-0.36)	0.34 (0.75-0.05)	.897

Note. IF: Item Facility. ID: Item Discrimination. FMC: Form Multiple-Choice. MMC: Meaning Multiple-Choice. IMP: Immediate Posttest. DP: Delayed Posttest

This study had a counterbalanced within-subject design to control for the effect of pictorial elucidation and etymological notes. All participants were exposed to the 24 idioms presented under four different conditions listed below (2). An example of each can be found in Appendix C.

2)

A) Condition 1: definition of idiomatic meaning;

B) Condition 2: definition of idiomatic meaning + picture;

C) Condition 3: definition of idiomatic meaning + etymological note;

D) Condition 4: definition of idiomatic meaning + picture + etymological note.

As shown in Table 3, participants were divided into four groups and randomly presented with four sets of six idioms under each of the four conditions aforementioned.

Table 3

Randomized presentation of idioms

	Group 1	Group 2	Group 3	Group 4
Idioms 1-6	A	B	C	D
Idioms 7-12	B	C	D	A
Idioms 13- 18	C	D	A	B
Idioms 19-24	D	A	B	C

Procedure

Figure 1 shows the overall procedure. All participants completed the pretest, the learning phase, and the immediate posttest over the course of an hour. First, for the pretest, participants answered a Qualtrics survey in which they solved the sentence completion task, the receptive form multiple-choice test, the definition task, and the receptive meaning multiple-choice test. The pretest might have had a learning effect, yet it serves to determine participants' prior knowledge and the degree of learning per idiom after the treatment.

Correct responses from the pretest will indicate what idioms the participants already knew and to what extent they were able to recognize and produce the meaning and form of those idioms before and after the treatment. Second, participants engaged in the learning phase, which involved using the "Flashcard" activity from Quizlet. They read the form of the idiom on the front side of a virtual card, and after clicking the card to flip it over, they had access to the meaning of the idiom. Third, participants solved an immediate posttest that consisted of the same Qualtrics survey with the same tests used in the pre-test. By the end of the session, all groups had been randomly exposed to all idioms in four different treatments. Finally, thirty- nine

participants had a delayed posttest one week later with the same tests used in the pretest and the immediate posttest, followed by a questionnaire about their background information (gender, age, first language, proficiency level, and international language exams). The number of participants that attended the second session from each institution was lower than in the first session (See Table 4), yet the number of participants per condition in each session was very similar (See Table 5). The delayed posttest for institution 4 was administered online to reduce attrition.

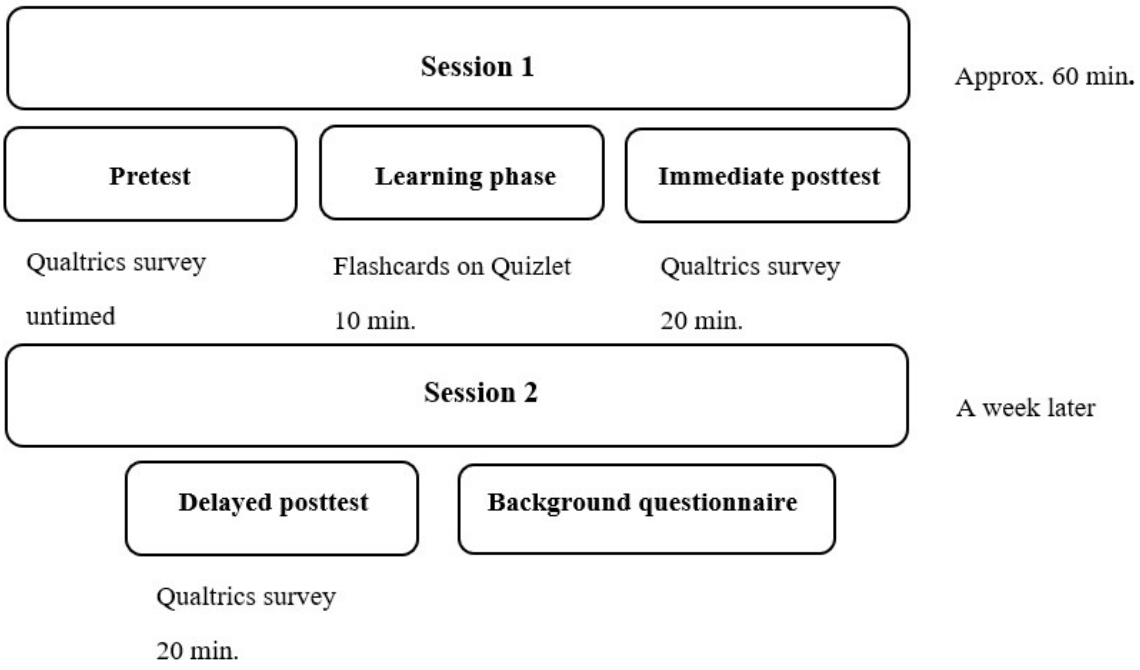


Figure 1. Procedure. Illustration of the steps taken in the implementation of this research project.

Coding and scoring

It should be acknowledged that as the data was being coded, a few instances of missing answers were identified in the multiple-choice tasks in the pretest and delayed posttest. The total

number of missing values was seven, one in the form multiple-choice task and six in the meaning the pretest and one in the delayed posttest. Idioms 18, 20, and 21 were missing only once each in

Table 4

Total number of participants per session and institution

	Session 1		Session 2	
	n	mode	*n*	mode
Institution 1	32	in-person	29	in-person
Institution 2	7	in-person	3	in-person
Institution 3	2	in-person	---	---
Institution 4	11	in-person	7	online
Institution 5	19	in-person	---	---
Total	71		39	

Note. *n* = number of participants.

Table 5

Total number of participants per session and per condition

	Condition 1	Condition 2	Condition 3	Condition 4	Total
Session 1	19	18	17	17	71
Session 2	8	9	11	11	39

Note. Participants over 30 (*n* = 4) were evenly distributed across conditions to avoid any age effect.

multiple-choice task. The answers for idiom 22 were missing in four instances, three of which in the delayed posttest. It is not clear whether the survey did not display the task items to the participants or whether it allowed students to skip the items. The missing data were considered as incorrect answers in the analysis.

The scoring of the answers of the multiple-choice tasks consisted of correct and incorrect answers. Sentence completion task and the definition task required different criteria. The scoring of the sentence completion task was performed following the criteria described below:

1. incorrect: 0 points;
2. partially correct: 1 point; this category includes answers with errors such as synonym substitution (e.g. *pink sheet*), incomplete idiom, i.e. one content word of two-word idioms or three-word idioms is missing (e.g. *ax to*…), incorrect content word, i.e. one content word of a two-word idiom or two words of a three-word idiom is/are correct and the other is incorrect (e.g. *walking duck*, *red herring* instead of *red tape* and vice versa, *take your gloves off*), number errors (e.g. *sour grape*), spelling errors (e.g. *loose canon*, except for *ax/axe to grind*), word order errors (e.g. *getting feet cold*), mistakes on function words (e.g. *pull some strings*, *bite his bullet*)
3. correct: 2 points; answers are considered correct regardless of tense choice.

The scoring for the answers of the definition task was performed according to the following criteria:

1. incorrect: 0 points; this category includes either a description of the individual literal meaning of the words that constitute the idiom (e.g. *last straw* defined as *the last thing*), or a definition that refers to a different idiom, words, or concept (e.g. *sour grapes* defined as *feeling angry*);
2. partially correct: 1 point; this category includes errors such as a vague definition with missing features of the concept (e.g. *green thumb* as *being good at something*), or an ambiguous concept (e.g. *dark horse* as *an unexpected person*);
3. fully correct: 2 points; the complete meaning of the idiom is expressed clearly; the definition has some spelling or grammatical errors, yet they do not affect the overall definition,

and therefore, the meaning of the idiom is fully conveyed (e.g. *red tape* as *rule and norms that are unnecesary);* even when the original answer is in Spanish, the translated version is considered correct if the concept is clear and complete (e.g. *let your hair down* defined as *relajarse,* in English *to relax*).

Rater 1 scored all the data. A second rater scored 20 % of the answers from the sentence completion task and the definition task. A Spearman's rank-order correlation was run to determine the correlation between raters' scorings at sentence completion task and definition task, which showed an inter-rater agreement of 99.5 % and 97.8 % accordingly.

Chapter 3 Results

Descriptive statistics

Table 6 shows the mean, the standard deviation, the range of answers, and the highest possible answer given in all tasks at pretest, immediate posttest, and delayed posttest. Values used for the receptive and productive tests vary based on the different total scores possible. Productive tests scores correspond to the addition of the values assigned to all the answers -- including fully correct answers, partially correct answers, and incorrect answers-- for each individual participant while the receptive tests scores resulted from the total number of correct answers each participant gave. Results show that participants performed best on the immediate posttest across all tasks. In productive tests, participants performed better on the definition task than on the sentence completion task at the pretest, immediate posttest, and delayed posttest; conversely, on the receptive tasks, participants performed better on the form multiple-choice tasks on the immediate and delayed posttest, except for the pretest, on which participants scored more highly for meaning than form.

Table 6

Mean, standard deviation, range of correct answers, and highest possible score

Test type	Task type	Phase								
		Pretest (*n* = 71)			Immediate posttest (*n* = 71)			Delayed posttest (*n* = 39)		
		M (*SD*)	Range	HPS	M (*SD*)	Range	HPS	M (*SD*)	Range	HPS
Productive	SCT	0.88 (2.80)	0−21	48	22.66 (11.91)	0−47	48	22.20 (14.19)	0−47	48
	Definition task	**5.86 (4.75)**	0−22	48	**24.19 (11.86)**	0−46	48	**22.64 (12.53)**	0−42	48
Receptive	Form MC	6.57 (3.60)	1−16	24	**22.50 (2.51)**	12−24	24	**21.65 (4.08)**	4−24	24
	Meaning MC	**8.39 (5.25)**	0−22	24	18.01 (5.50)	0−24	24	17.84 (5.40)	0−24	24

*Note***.** SCT = sentence completion task. MC = multiple choice. HPS = highest possible score.

Inferential statistics

The number of correct and incorrect answers across tasks and conditions are displayed at immediate and delayed time points (See Figures 2, 3, 4 & 5). Condition 1 corresponds to the definition of the idiomatic meaning of the idioms alone, condition 2 is the definition and the picture, condition 3 is the definition and the etymological explanation, and condition 4 combines the definition, the picture, and the etymological explanation. Figure 2 indicates the percentage of correct, partially correct, and incorrect answers for the productive tests, sentence completion and definition task, across conditions at the immediate posttest. The number of correct answers for condition 1, 2 and 4 was not higher than 45 % in both productive tasks. However, condition 3 showed the lowest number of correct answers in sentence completion task. The number of incorrect answers was as high as 20 % in all conditions, except for condition 3 at sentence completion and definition tasks that reached up to 50 % and 45 % respectively. The number of partially correct answers remained below 20%, except for condition 1 and 2 in the definition task that slightly exceeded that percentage.

Figure 3 shows the percentage of correct, partially correct, and incorrect answers for the productive tests across conditions at the delayed posttest. Condition 2 and 4 showed the highest number of correct answers in sentence completion task (45 %), while condition 2, 3 and 4 scored the highest number of correct answers in the definition task (almost 20 %). The highest number of incorrect answers in the sentence completion task and definition task was for condition 3 again (50 % and 45 % respectively). The number of partially correct answers was similar across conditions in each task. The percentage is slightly higher for the conditions in the definition task than for the sentence completion task. The highest number of partially correct answers was for condition 1 in the definition task.

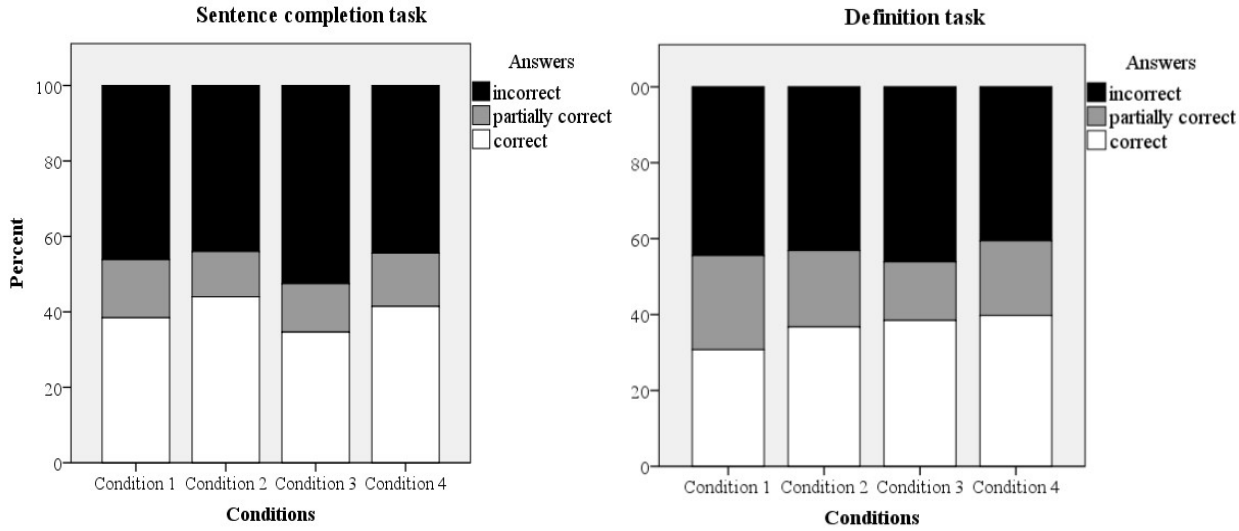


Figure 2. Percentage of correct and incorrect answers at the immediate posttest.

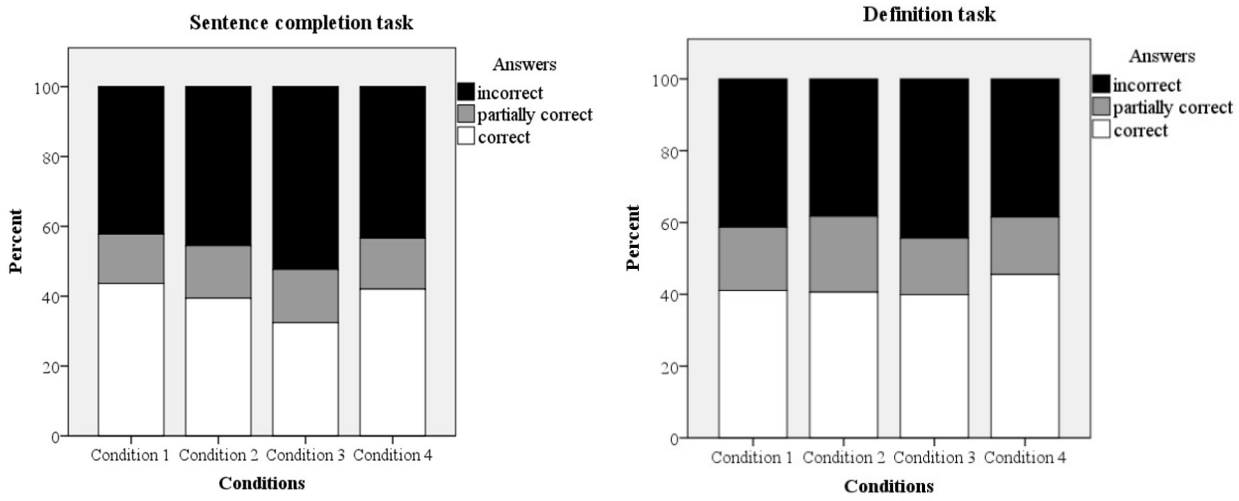


Figure 3. Percentage of correct and incorrect answers at the delayed posttest

Figure 4 displays the percentage of correct and incorrect answers scored per condition in the form and meaning multiple-choice tasks at the immediate posttest. In the receptive form

multiple-choice task, the percentage of correct answers is above 90% in all conditions while the percentage of correct answers at the meaning multiple-choice task is above 70% in each

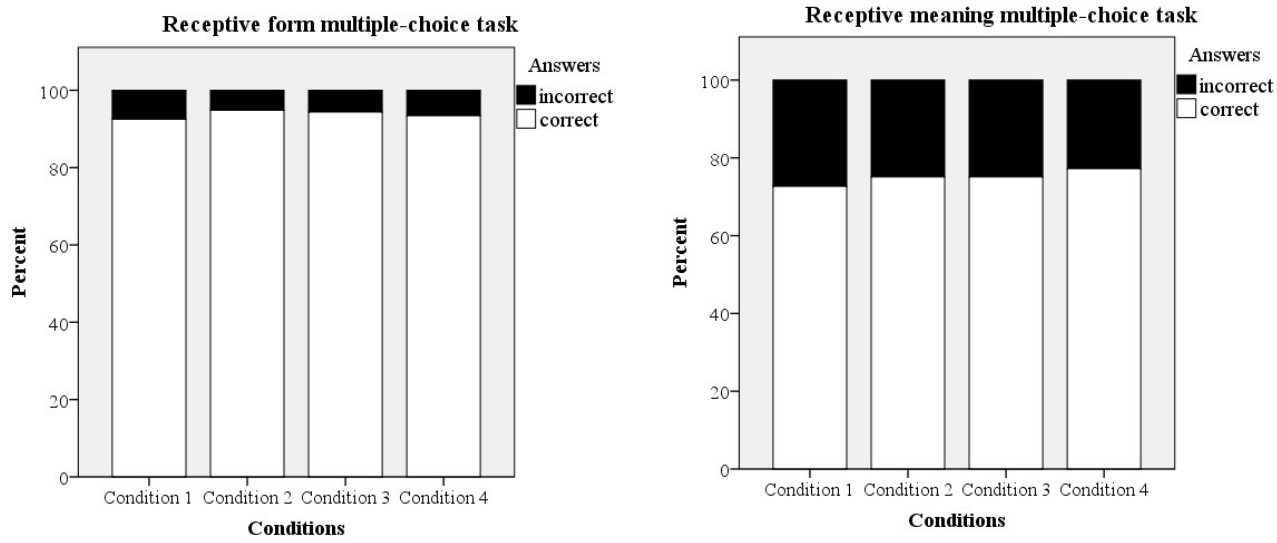


Figure 4. Percentage of correct and incorrect answers at the immediate posttest

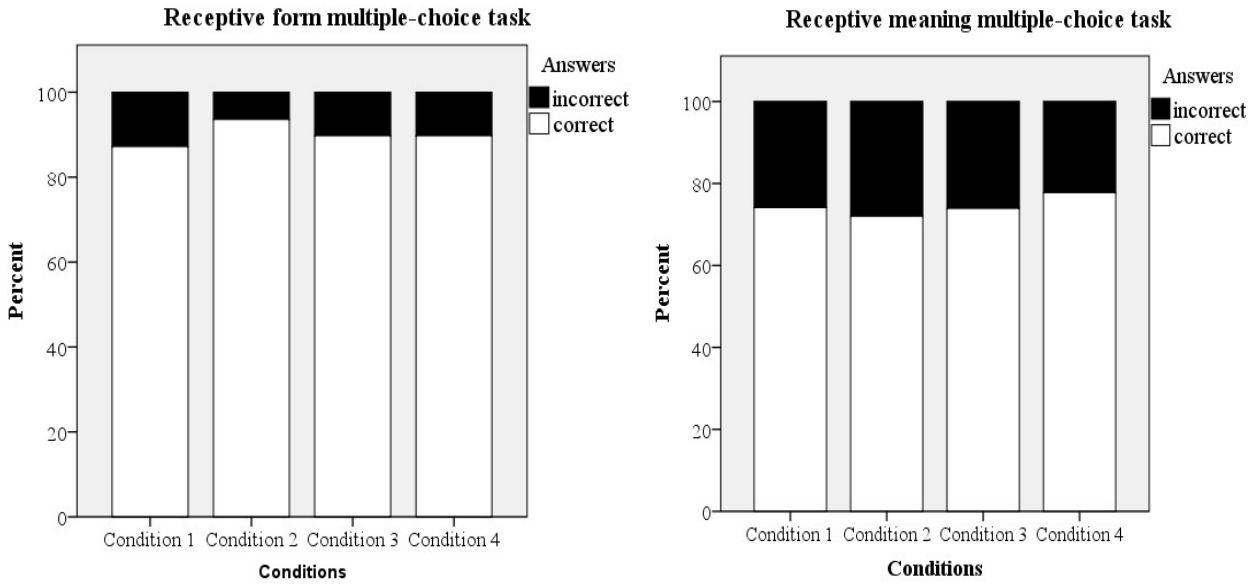


Figure 5. Percentage of correct and incorrect answers at the delayed posttest

condition. These results show a ceiling effect in both multiple-choice tasks. The percentage of correct answers and incorrect answers did not vary across conditions in the immediate posttest. Figure 6 shows the percentage of correct and incorrect answers for the same tasks in each condition at the delayed posttest. The graphs indicate that both receptive tasks in the delayed posttest had similar results across conditions to those in the immediate posttest.

Table 7 shows mean, standard deviation, Friedman's and Wilcoxon-signed rank's tests across tests, tasks, and conditions. Participants were exposed to four different conditions in all tasks and time points. In the productive tasks, the highest possible mean score for each individual condition was a total of 12, while the highest possible mean score per condition in the receptive tasks was 6. Since the assumption of normality was not met, nonparametric statistical analyses were performed accordingly. Friedman tests were conducted to determine what conditions had a significant effect in the productive tests, sentence completion and definition tasks, at immediate and delayed time point. The analysis was based on the addition of the scores given for correct answers. Results showed no statistically significant differences in the definition task at immediate posttest ($\chi 2$ (3) = 4.04, p = .258) and delayed posttest ($\chi 2$ (3) = 2.63, p = .453), and in sentence completion task at delayed posttest ($\chi 2$ (3) = 10.14, p = .150).

However, Friedman tests showed there was a statistically significant difference in sentence completion task at immediate posttest across conditions at the 5% level of significance ($\chi 2$ (3) = 12.47, p = .006). The corresponding post-hoc pairwise Wilcoxon singed rank tests with a Bonferroni correction applied (i.e., the adjusted significance level is 0.05/6 = 0.0083) indicated there were statistically significant differences between condition 1 (definition) and condition 3 (definition and etymological note), and between condition 3 (definition and etymological note) and condition 4 (definition, picture, and etymological note).

Table 7

Mean, standard deviation, Friedman's and Wilcoxon-signed rank's tests across tests, tasks, and conditions

Test type	Task type	Time point	C	Mean (SD)	Median	Friedman test *p*-value	Wilcoxon-signed rank test *p*-value (effect size*)		
						C1--C4	C2	C3	C4
P	STC	IMM	C1	**6.08** (3.83)	7.00	**.006**	.316 (0.08)	**.004 (0.24)**	.495 (0.06)
			C2	5.73 (3.28)	6.00			.034 (0.03)	.491 (0.06)
			C3	4.93 (3.48)	4.00				**.001 (0.27)**
			C4	**5.92** (3.56)	6.00				
		D	C1	5.54 (3.82)	4.00	.150	.236 (0.13)	.199 (0.15)	.602 (0.06)
			C2	**6** (4.00)	4.00			.038 (**0.24**)	.562 (0.07)
			C3	4.92 (3.69)	4.00				.013 (**0.28**)
			C4	**5.74** (3.97)	5.00				
	DEF	IMM	C1	5.94 (3.51)	6.00	.258	.652 (0.09)	.509 (0.03)	.125 (0.16)
			C2	**6.08** (3.51)	6.00			.446 (0.03)	.197 (0.10)
			C3	5.75 (3.31)	6.00				.037 (0.14)
			C4	**6.42** (3.45)	6.00				
		D	C1	5.36 (3.38)	5.00	.453	.337 (0.12)	.574 (0.12)	.107 (**0.22**)
			C2	**5.74** (3.57)	6.00			.781 (0.00)	.456 (0.10)
			C3	5.54 (3.39)	5.00				.255 (0.13)
			C4	**6** (3.53)	6.00				
R	F	IMM	C1	5.55 (0.97)	6.00	.691	.070 (0.15)	.485 (0.06)	.763 (0.03)
			C2	**5.69** (0.77)	6.00			.632 (0.04)	.413 (0.07)
			C3	**5.66** (0.81)	6.00				.487 (0.06)
			C4	5.61 (0.84)	6.00				
		D	C1	5.23 (1.39)	6.00	.108	.012 (**0.28**)	.307 (0.12)	.193 (0.15)
			C2	**5.62** (0.88)	6.00			.177 (0.18)	.154 (0.16)
			C3	5.38 (1.04)	6.00				.971 (0.00)
			C4	5.38 (1.25)	6.00				
	M	IMM	C1	4.37 (1.69)	5.00	.278	.610 (0.04)	.461 (0.06)	.040 (0.17)
			C2	4.46 (1.63)	5.00			.803 (0.02)	.132 (0.13)
			C3	**4.51** (1.54)	5.00				.151 (0.12)
			C4	**4.68** (1.54)	5.00				
		D	C1	4.41 (1.74)	5.00	.147	.572 (0.06)	.791 (0.03)	.191 (0.15)
			C2	4.31 (1.58)	5.00			.537 (0.07)	.086 (**0.20**)
			C3	**4.44** (1.55)	5.00				.278 (0.12)
			C4	**4.67** (1.47)	5.00				

Note. P = productive. R = receptive. STC = sentence completion. DEF = definition. F = form. M= meaning. IMM = immediate. D = delayed. C = condition type. *effect size measures of Wilcoxon signed rank test.

At the Bonferroni adjusted significance level, there was no statistically significant difference between condition 2 (definition and picture) and condition 3, yet the corresponding raw p-value was smaller than 0.05, the unadjusted level of significance. In remaining comparisons in sentence task at immediate posttest, even raw p-values showed insignificant results. Condition 3 had the lowest average scores in sentence task at immediate posttest. However, the p-value of sentence completion task is insignificant for condition 3 at delayed posttest.

Again, Friedman tests were conducted at the immediate and delayed posttest respectively in order to see whether participants' answers of the form and/or meaning of idioms from the receptive tests were significantly different across conditions. None of the results from the receptive test showed significant differences across conditions at any time point at the .05 significance level. There was no statistically significant difference in the form of idioms independently from the condition used in the learning phase at either immediate ($\chi 2$ (3) = 1.46, p = 0.691) or delayed time point ($\chi 2$ (3) = 6.08, p = 0.108). Likewise, Friedman tests also showed that there was no statistically significant difference for the meaning of idioms regardless of the type of condition used in the learning phase at either immediate ($\chi 2$ (3) = 3.85, p = 0.278) or delayed time point ($\chi 2$ (3) = 5.37, p = 0.147).

Table 7 also shows effect sizes at Wilcoxon signed rank test across productive and receptive tasks, conditions, and immediate and delayed time points. The effect sizes were small on all tasks, conditions, and time combinations. The conditions that showed a statistically significant difference in the sentence completion task at immediate posttest also had small effect sizes. (condition 1 and condition 3, d = 0.24, and condition 3 and condition 4, d = 0.27).

The last analysis was performed on the learning impact of the treatment on individual idioms. This analysis was made to know whether participants' knowledge of each idiom significantly increased from the pretest to the immediate posttest. The analysis also aimed at identifying idioms whose form and meaning might be more susceptible to be learned receptively and/or productively than others. To examine the intervention effect on individual idioms at definition and sentence completion tasks, Marginal Homogeneity tests (nonparametric significance test for multiple variables) were used for each given idiom. Results showed that there was a statistically significant difference between the answers from the pretest and the immediate posttest at each idiom. In other words, the intervention effect was statistically significant, indicating high learning gains for each individual idiom.

McNemar's tests (nonparametric significance test for dichotomous variables) were used to examine the treatment effect on each specific idiom in both receptive tasks. This test was applied to two proportions of correct answers given per idiom, one being the number of correct answers at the pre-test, and the other the number of correct answers at the immediate posttest. Bonferroni adjusted alpha levels of .0021 per test were used for the analyses of form and meaning respectively. Results displayed statistically significant differences between pretest and immediate posttest at each idiom in the form multiple-choice task, and meaning multiple-choice task, except for *let your hair down*, *last straw*, *pull the strings*, and *dark horse* in the meaning multiple-choice task.

Table 8 shows the number and the percentage of participants who gave correct answers for each idiom in the form multiple-choice task in both the pretest and immediate posttest. The idioms with the lowest number of correct answers in the pretest were *grass roots* (n = 2, 3%), *lame duck* (n = 5, 7%), *lock horns* (n = 8, 11%), and *pink slip* (n = 8, 11%). The idioms with the

highest number of correct answers in the pretest were *getting cold feet* (n=48, 68 %), *the gloves are off* (*n* = 42, 59%), and *pull the strings* (*n* = 37, 52 %). After the treatment, the idioms that showed the highest improvement were *grass roots* (31 %), *lame duck* (11%), *pink slip* (8%), *lock horns* (7%), *red herring* (5%), *loose cannon* (5%), and *sacred cow* (5%).

Table 8

Number and percentage of participants with correct answers for each idiom in the receptive form multiple-choice task in the pretest and immediate posttest

		N of participants who gave correct answers for each idiom (%)		McNemar's test
Idiom *N*	Idiom	Pretest	Immediate posttest	*p*-value
1	Sitting duck	14 (24 %)	64 (90 %)	< 0.001
2	Loose cannon	11 (16 %)	67 (94 %)	< 0.001
3	White elephant	23 (32 %)	70 (99 %)	< 0.001
4	Sour grapes	15 (21 %)	67 (94 %)	< 0.001
5	Grass roots	2 (3%)	68 (96 %)	< 0.001
6	Red herring	11 (16 %)	68 (96 %)	< 0.001
7	Lame duck	5 (7 %)	62 (87 %)	< 0.001
8	Red tape	16 (23 %)	70 (99 %)	< 0.001
9	Ax to grind	16 (23 %)	67 (94 %)	< 0.001
10	The gloves are off	42 (59 %)	69 (97 %)	< 0.001
11	Let your hair down	29 (41 %)	71 (100 %)	< 0.001
12	Lock your horns	8 (11 %)	64 (90 %)	< 0.001
13	Walk the plank	17 (24 %)	66 (93 %)	< 0.001
14	Getting cold feet	48 (68 %)	68 (96 %)	< 0.001
15	Move the goalposts	24 (34 %)	63 (89 %)	< 0.001
16	Bite the bullet	13 (18 %)	68 (96 %)	< 0.001
17	Sacred cow	11 (16 %)	64 (90 %)	< 0.001
18	Green thumb	26 (37 %)	69 (97 %)	< 0.001
19	Last Straw	33 (47 %)	64 (90 %)	< 0.001
20	Acid test	23 (32 %)	70 (99 %)	< 0.001
21	Pull the strings	37 (52 %)	60 (85 %)	< 0.001
22	Dark horse	17 (24 %)	67 (96 %)	< 0.001
23	Jump the gun	23 (32 %)	65 (92 %)	< 0.001
24	Pink slip	8 (11 %)	67 (94 %)	< 0.001

Note. N= number.

Table 9 displays the number and percentage of participants that provided the correct answers for each idiom in the receptive meaning multiple-choice task in the pretest and immediate posttest. The idioms with the lowest number of correct answers in the pretest were *grass roots* (*n* = 8, 11%), *red herring* (*n* = 9, 13%), and *ax to grind* (*n* = 11, 16%). The idioms

with the highest number of correct answers in the pretest were *let your hair down* (*n* = 60, 85%) and *getting cold feet* (*n* = 50, 70%). The idioms that had the highest percentage increase were *grass roots* (5%), *red herring* (4%), *red tape* (3%), *dark horse* (2%), *sitting duck* (2%), *sour grapes* (2%), *pink slip* (2%), and *loose cannon* (2%). Overall, results showed that the learning of

Table 9

Number and percentage of participants with correct answers of each idiom in the receptive meaning multiple-choice task in the pretest and immediate posttest

		N of participants who gave correct answers to each idiom (%)		McNemar's test
Idiom N	Idiom	Pretest	Immediate posttest	*p*-value
1	Sitting duck	22 (31 %)	62 (87 %)	< 0.001
2	Loose cannon	23 (32 %)	59 (83 %)	< 0.001
3	White elephant	29 (41 %)	63 (89 %)	< 0.001
4	Sour grapes	20 (28 %)	55 (78 %)	< 0.001
5	Grass roots	8 (11 %)	46 (65 %)	< 0.001
6	Red herring	9 (13 %)	47 (66 %)	< 0.001
7	Lame duck	13 (18 %)	38 (54 %)	< 0.001
8	Red tape	15 (21 %)	53 (75 %)	< 0.001
9	Ax to grind	11 (16 %)	48 (68 %)	< 0.001
10	The gloves are off	30 (42 %)	63 (89 %)	< 0.001
11	Let your hair down	60 (85 %)	69 (97 %)	0.022
12	Lock your horns	18 (25 %)	62 (87 %)	< 0.001
13	Walk the plank	32 (45 %)	50 (70 %)	< 0.001
14	Getting cold feet	50 (70 %)	64 (90 %)	0.001
15	Move the goalposts	24 (34 %)	49 (69 %)	< 0.001
16	Bite the bullet	36 (51 %)	57 (80 %)	< 0.001
17	Sacred cow	12 (17 %)	29 (41 %)	< 0.001
18	Green thumb	37 (52 %)	54 (76 %)	0.001
19	Last Straw	28 (39 %)	37 (52 %)	0.035
20	Acid test	22 (31 %)	52 (73 %)	< 0.001
21	Pull the strings	38 (54 %)	52 (73 %)	0.003
22	Dark horse	17 (24 %)	57 (81 %)	0.003
23	Jump the gun	28 (39 %)	59 (83 %)	< 0.001
24	Pink slip	20 (28 %)	53 (75 %)	< 0.001

Note. N = number.

the form of idioms was higher than the meaning, yet both were statistically significant.

A qualitative analysis of the partially correct answers indicated what error types were frequent at sentence completion tasks. Table 10 displays the use of synonyms and incorrect

content word as the first two most frequent. There were also other less frequent error types listed such as mistakes on function words, spelling mistakes, incomplete idioms, and incorrect idioms.

In the definition task, partially correct answers were also analyzed. Those idioms that were mostly defined using partial features or ambiguous concepts in the immediate posttest were *the gloves are off* (e.g. fighting), *walk the plank* (e.g. getting fired), *move the goalposts* (e.g. changing rules), *red tape* (e.g. reference to something unnecessary), acid test (e.g. an important test), and *white elephant* (e.g. something expensive). The only idiom that did not have partially correct answers was *let your hair down*. In the delayed posttest, those idioms with the highest number of partially correct answers were again the *gloves are off*, *walk the plank,* and *move the goalposts, and jump the gun* (e.g. to go ahead).

Table 10

Error types in sentence completion task at immediate and delayed posttests

Error type	Idiom	Examples
Synonyms	let your hair down	*lay your hair down; leave your hair down; lie your hair down.*
	bite the bullet	*beat the bullet; bit the bullet; bruise the bullet.*
	pink slip	*pink sheet; pink letter; pink paper.*
Incorrect content word	pink slip	*pink stick; pink card; pink face.*
	ax to grind	*ax to greed; ax on hand; ax to rid.*
	move the goalposts	*move the goalstand; move the goals; move the goalkeep.*
Function words	----	*bite your bullet; pull some strings; the gloves are on.*
Spelling mistakes	----	*seeting duck; loose canon; green thomb.*
Incomplete idioms	pink slip	*pink*
	red herring/ red tape	*red*
	ax to grind	*axe to*
Incorrect idiom	lame duck*	*loose cannon; last straw; lock horns.*
	lock horns*	*the gloves are off*
	red herring**	*red tape*
	lame duck, loose cannon, last straw***	*lock horns*

Note. **lame duck* and *lock horns* were mistakenly used in place of the above listed idioms. ***red tape* was also used for *red herring.* ****lock horns* was the idiom with the most various number of incorrect idioms.

Chapter 4 Discussion

The first two research questions examined the effect of pictures on the form and meaning recognition and recall of idioms in the immediate and delayed posttests. The condition with pictures was not statistically significant at any task type nor any time point. However, this result does not necessarily indicate that pictures have no effect, or even further, a negative effect on the learning of idioms. This condition yielded the highest mean scores across conditions in the sentence completion task at delayed posttest, and in the receptive form multiple-choice tasks at immediate and delayed posttests. These results might suggest that pictures might have the potential to facilitate structural elaboration (i.e. form knowledge of idioms). In other words, pictures might have a stronger facilitative effect on form than meaning in both productive and receptive tests. This apparent facilitative effect of pictures clashes with Boers' (2009) claim that pictures have a detrimental effect on the learning of idioms, yet supports Szczepaniak and Lew's (2011) study that showed that the retention of form of idioms can be significantly stimulated by pictures in the immediate and delayed time points.

The condition with pictures also had high mean scores on the definition task at immediate and delayed posttests. This could be an indicator of the potential pictures might have to contribute to promoting productive knowledge of the meaning of idioms. Additional research should be done to further explore this tendency since the focus of previous research has solely been on meaning recognition using multiple-choice tasks, rather than using a definition task to assess meaning production as well.

Though previous studies argue that pictures help to enhance the meaning of idioms in meaning receptive tests (Boers, 2009; Szczepaniak and Lew, 2011), this study failed to support this claim. Condition 4 (picture and the etymological notes combined) scored the highest mean

scores in the receptive meaning multiple-choice task, followed by condition 1 (definition). Similarly, condition 4 and condition 1 were the conditions that were most successful in stimulating form productive knowledge through the sentence completion task at immediate time point. It can be speculated that condition 4 or 1 could have had a cancellation effect. Some elements could have drawn participants' attention more than others; the definition element that was given for all idioms in all conditions might have been one. It should be acknowledged that this study did not control the amount of time participants spent on each individual flashcard during the learning phase. Future studies should investigate whether the same results can be obtained even if selective attention and time are controlled. These problems could be addressed by using another resource in the learning phase. For example, one as simple as a video could be an alternative to present elements separately and control for the time participants spend on the information provided per idiom. Such an instrument could ensure all participants are exposed to the same information for the same amount of time.

The third and fourth research questions studied whether etymological notes enhanced the recognition and recall of form and meaning of idioms in the immediate and delayed posttest. Like the condition with pictures, the condition with etymological notes alone (condition 3) was not statistically significant at any task nor any time point. In fact, it accounted for the lowest mean scores in sentence completion task and definition task at immediate posttest, displaying the highest percentage of incorrect answers. A statistically significant difference between this condition and condition 1, and condition 4 was observed only in sentence completion task at immediate posttest. This might be evidence that etymological notes may have a negative effect on idiom learning in the form production of idioms while other conditions seem to be more effective in stimulating learning than etymological explanations. Etymological notes might

interfere with the structural elaboration of idioms (form knowledge of idioms). This finding is in line with Szczepaniak and Lew's (2011) study in which etymological notes fostered neither form nor meaning recognition and retention of idioms. However, in this study, the negative effect of condition 3 seems to lessen over time, suggesting that etymological notes might not necessarily be detrimental for idiom learning in the long-run.

Like productive tests, receptive tests did not indicate any statistically significant difference between the condition 3 and the other conditions. Still, mean scores were higher for etymological note condition than for picture condition in the receptive meaning multiple-choice tasks at immediate and delayed posttests. This tendency is congruent with Boers (2008) that claimed that etymological notes favored meaning recognition of idioms at immediate time point.

The analysis of individual idioms showed that the treatments were effective. There were significant learning gains after the treatment in most idioms in all tasks. Interestingly, participants showed higher learning of the form of idioms than the meaning in the multiple-choice tasks. This might be attributable to the design of the study. Participants were exposed to the form of idioms 12 times while they were presented with the definition of idioms just 4 times, including the treatment and all tasks at all time points. Therefore, it is not surprising if participants were able to recognize the form of idioms better than the meaning. Folse (2004) affirms that the frequency of exposure of new vocabulary conditions vocabulary learning. Every time participants had to solve one of the tasks, they were repeatedly exposed to the form of the idioms. Through this repeated exposure, participants might have been encouraged to notice the vocabulary (Folse, 2004), which is likely to have drawn students' attention to the words and stimulated form recognition of the idioms.

Some idioms such as *grass roots, red herring, pink slip*, and *loose cannon showed* high learning gains in both receptive tasks. Very few idioms showed no significant learning gains in the receptive meaning multiple-choice task. One was *let your hair down*, which most participants already displayed meaning receptive knowledge of before the treatment, showing that students at upper-intermediate and low-advanced level are already able to recognize the meaning of this idiom. Other idioms such as *last straw, pull the strings*, and *dark horse* also had no significant learning improvement in the same task after the treatment, which revealed low learnability among this population.

The analysis of partially correct answers in productive tasks also showed that idioms with more than two words such as *ax to grind* and *move the goalposts* displayed multiple erroneous forms. Another source of difficulty were idioms whose words started with the same letter (e.g. *lame duck*, *loose cannon*, *last straw*, and *lock horns*). Additionally, idioms with the same words (e.g. *red herring* and *red tape)* or similar meaning (*the gloves are off* and *lock horns*) accounted for several instances of partially correct answers. Participants confused these idioms and used them interchangeably at times. As indicated by Folse (2004), thematic learning has proven to be more useful than semantic clustering of new L2 vocabulary, the latter being the one that usually interferes with learning. Future research should consider selecting words that have neither a similar structure nor semantic domain to avoid confusion and mismatch between concepts.

The selection criteria of idioms and pictures could still be further explored. In this study, idioms consisting of an adjective plus a noun might have been more illustrable than idioms made of a verb and a complement. Nouns are more concrete in nature than actions, which could have been an asset for the learning of the form of these idioms. Thus, the effect of pictures on idiom

learning might also depend on the potential of the expression to be illustrated by a picture. Additionally, it should be acknowledged that in this study the pictures selected for *the gloves are off*, *move the goalposts*, and *walk the plank* could have also illustrated some metaphorical meaning. Future studies should use pictures that illustrate the literal meaning of individual words without expressing any metaphorical meaning of the idioms.

To sum up, all conditions were effective and the differences between them were not statistically significant, except for condition 1 and 3, and condition 3 and 4 in the sentence completion task at immediate posttest. Mean values were used as indicators of the potential that pictures seem to have of facilitating knowledge of form of idioms, especially in the productive recall and recognition of idioms in the long run. On the other hand, mean scores seem to show that etymological notes were the least effective in the production of idioms, indicating that etymological explanations might have a detrimental effect on the production of the form and the meaning of idioms in the short-term. Further research needs to be done to provide more evidence on the facilitative effect of pictures and the apparent detrimental impact of etymological notes on the learning of idioms.

As for the implications of this study, they are two-fold. Far from having a negative impact on learning, pictures can foster the learning of new idioms across tasks and time. Learners can benefit from pictures – that are clear, unambiguous, and reflect the literal meaning of the words –to learn new idioms. Another implication is that etymological notes may posit some difficulty if used in the first encounters students have with the new vocabulary. It seems that students might not be ready to process such verbal explanation of the origin of the idioms at that stage, and it is likely that etymological notes cause confusion, rather than stimulating the

learning of idioms, especially the form. Hence, teachers should use them with caution to prevent fuzziness and erroneous conceptualizations.

Further research needs to be done using a similar design to this study. Previous research has only investigated the meaning recognition of idioms and form production of idioms. Previous studies have not drawn much attention on form recognition and meaning production of idioms, which should be used to make vocabulary analysis of idioms more balanced and understand all aspects of vocabulary learning. It would be interesting to use alternate instruments that might better assess the receptive and productive knowledge of idioms. For example, instead of using a form multiple-choice task, a gap-fill activity with a vocabulary box could be administered to assess participants' ability to recognize the form and the meaning of idioms.

Future studies could analyze other factors not discussed here. Studies on other L2 populations with a large sample size could contribute to providing substantial evidence of the conditions that may enhance or hamper vocabulary learning. Controlling the time of exposure to the new vocabulary is another factor that can provide more information about the effectiveness of some conditions over the others. Future studies should also consider administering the delayed posttest at least two weeks after the treatment to explore retention over an extended period of time. This study provides a foundation for subsequent research oriented to unveil the complexity of idiom learning.

APPENDICES

APPENDIX A: Idioms, definitions, etymological notes, and pictures

1) Sitting duck

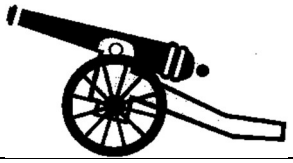
Definition: If you refer to someone as a sitting duck, you mean that they are an obvious target, and that it is very easy to attack them or criticize them.

Etymological note: A duck is an easy target for hunters when it is sitting on the water or on the ground.

2) Loose cannon

Definition: If you describe someone as a loose cannon, you mean that their behavior is unpredictable and therefore could have unfortunate or dangerous consequences.

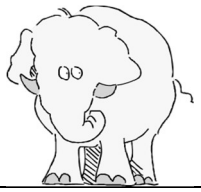
Etymological note: This expression refers to the cannons which used to be carried on the decks of warships. If one of the cannons was not properly fastened down, it could spin round and make a hole in the ship.



3) White elephant

Definition: If you describe something such as a new building, plan, or project as a white elephant, you mean that it is a waste of money and completely useless.

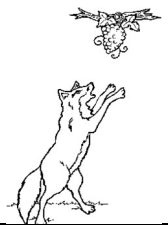
Etymological note: The origin of the expression is a traditional custom from Siam, present-day Thailand. If a rare albino (white) elephant was captured, it was the property of the emperor, and only he could ride or use the animal. Whenever the emperor wished to ruin someone who displeased him, he would give the man a white elephant. The man would then be forced to feed and care for the animal but could neither use nor destroy it.



4) Sour grapes

Definition: If you describe someone's attitude or behavior as sour grapes, you mean that they are pretending that something they cannot have is of little value or interest.

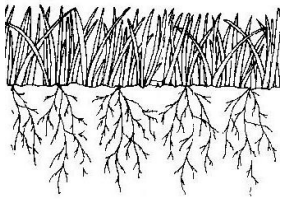
Etymological note: The expression originates from Aesop's fable about a fox that wanted some grapes from a vine but was unable to jump high enough to get them. In disgust, he claimed that they weren't worth having anyway because they were probably sour.



5) Grass roots

Definition: The grass roots of an organization or movement are the ordinary people who form the main part of it, rather than its leaders.

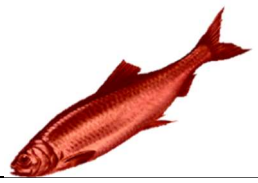
Etymological note: The expression used to mean "that which is rooted in the earth, that is, that which has its origins among the peasantry (=farmers) or common folk" (=people in general).



6) Red herring

Definition: If you describe a piece of information, a suggestion, or an action as a red herring, you mean that it is irrelevant and, often deliberately, is taking people's attention away from the main subject, problem, or situation that they should be considering.

Etymological note: The expression originates from the fact that a red (smoked) herring has a powerful and persistent odor. Centuries ago red herring was used to train dogs to track scents. Those hoping to misdirect tracking dogs would drag red herring across their trails since a dog that gets a smell of red herring will lose any other scent that it has been following.



7) Lame duck

Definition: If you say that someone or something is a lame duck, you are criticizing them for being in a very weak position and in need of support.

Etymological note: The expression suggests that a lame duck—a duck that cannot fly—is ineffectual (= weak). It originally comes from the 1760s London Stock Market, where it referred to investors who were unable to pay their debts.

8) Red tape

Definition: People refer to official rules, unnecessary paperwork and procedures as red tape when they seem unnecessary and cause delay.

Etymological note: Lawyers and government officials used to tie documents together with red or pink tape.

9) Ax to grind

Definition: If someone has an axe to grind, they have private, often selfish reasons for being involved in something.

Etymological note: There are several explanations for the origin of this expression. One is a story told by Benjamin Franklin about a man who managed to get his own axe sharpened by asking a boy to show him how his father's grindstone (= a round stone to make tools sharp) worked.

10) The gloves are off

Definition: If you are talking about a situation in which people have decided to fight or compete aggressively with each other, you can say the gloves are off.

Etymological note: The reference here is to boxers fighting with bare fists, which is more dangerous than fighting with gloves on.

11) Let your hair down

Definition: If you let your hair down, you relax and enjoy yourself, and do not worry about being dignified or behaving correctly.

Etymological note: In the past, women wore their hair tied up when they were in public and only untied their hair in private or in informal situations.

12) Lock your horns with sb

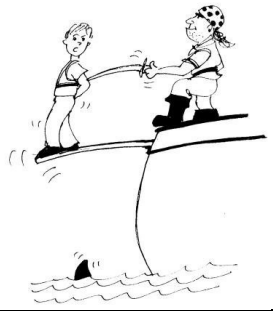
Definition: If you lock horns with someone, you argue or fight with them.

Etymological note: The reference here is to two male animals, such as deer, fighting over a female and getting their horns caught together or 'locked'.

1**3) Walk the plank**

Definition: If something goes wrong and someone in a position of authority walks the plank, they have to accept responsibility for what has happened and leave their position.

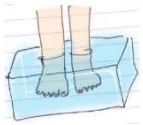
Etymological note: Many people believe that pirates used to kill their prisoners by forcing them to walk along a plank sticking out from the edge of a ship until they fell into the sea.



14) Get cold feet

Definition: If you get cold feet or have cold feet about something, you are not sure whether you want to do it, or you become too nervous and worried to do it.

Etymological note: The expression originates in a belief that when someone is suddenly scared, their circulation is poor and the blood leaves their feet and hands, which, as a result, turn very cold.



15) Move the goalposts

Definition: If you accuse someone of moving the goalposts, you mean that they have changed the rules, policies, or aims in a situation or activity, in order to gain an advantage for themselves and to make things more difficult for the other people involved.

Etymological note: A goalpost is one of the two upright wooden posts that are connected by a crossbar and form the goal in games such as football and rugby. If you move the goalpost, you make things more difficult for other players.

16) Bite the bullet

Definition: If you bite the bullet, you face a difficult or unpleasant situation.

Etymological note: the expression originates from the practice where, before the days of anesthesia, a person undergoing an operation might have been told to bite down on a bullet to distract from the pain.

17) Sacred cow

Definition: If an idea or a thing is a sacred cow, it cannot be altered.

Etymological note: The expression originates from the Hindu belief that cows are sacred and cannot be killed or eaten. The expression is frequently used to describe an idea or thing that ought to be changed or altered but cannot be because the authority forbids it.

18) Green thumb

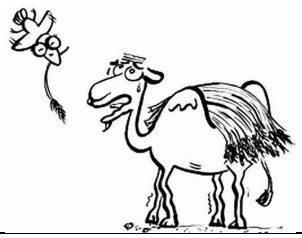
Definition: if you have a green thumb, you are good at making plants grow.

Etymological note: The expression suggests that success with growing plants is a result of having a thumb that is the color of healthy plants.

19) Last straw

Definition: If something is the last straw, it is the final thing, or the thing or action that is too much or goes too far.

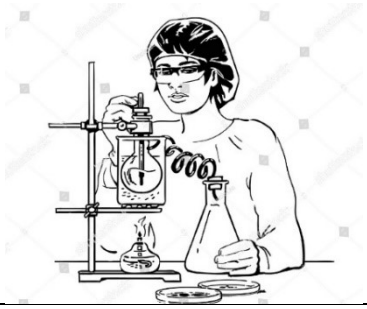
Etymological note: This expression suggests the idea of loading straw (a relatively light material) onto a camel's back until one final light straw (the last straw) breaks the camel's back.



20) Acid test

Definition: If something is a test acid, it is the most important or crucial test of worth; a way of deciding whether something is successful or true.

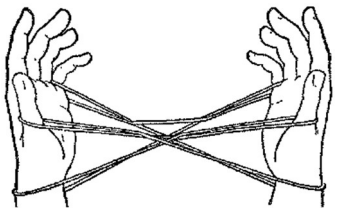
Etymological note: The expression originates from the use of nitric acid on gold to determine whether the gold was genuine.



21) Pull strings

Definition: If someone pulls strings, they use their influence to control events or the actions of other people.

Etymological note: The expression originates from the idea of a string puppet or marionette, which can be controlled by pulling on its strings. A person who can pull strings can control a situation and influence others.



22) Dark horse

Definition: If a person taking part in a race, etc. surprises everyone by winning, he/she is a dark horse.

Etymological note: The expression originates from horse racing jargon. It is often used to mean a surprise candidate in a political election.

23) Jump the gun

Definition: If you jump the gun, you do something prematurely, or you start early, before all the preparations have been made.

Etymological note: The expression probably originates from foot racing, in which an overly anxious runner would accidentally begin the race before the starting gun was fired.

24) Pink slip

Definition: If you get a pink slip, you are fired from your job.

Etymological note: The expression probably originates from the color of the form used to notify people that they had been fired. Such forms often came in multiple carbon copies. Each copy was a different color and was designated for a different recipient, e.g., the pay office got one particular color, while the fired person always got the pink copy.

APPENDIX B: Research Instruments

Productive form recall test

In this first activity, you will fill out the blank spaces with the correct idiom. You have a clue for the first letter of the idiomatic expression. You do NOT need to guess any answer. If you do not know the answer, write an "X" in the box.

1) Unarmed policemen walking the streets late at night are ***s***___ ___.
2) He has a reputation as a ***l***___ ___ whose comments sometimes upset Wall Street.
3) The salesman has been trying to get rid of that car for more than a year. It costs too much to run and insure, so no one wants it—it's a ***w***___ ____
4) When she failed the entrance exam, she started saying that she never wanted to go to college anyway, but I think that's just ***s***___ ___.
5) The feeling among the ***g***___ ___of the party is that the leaders aren't radical enough.
6) This is a total political ***r***___ ___ and an attempt to divert from the main issues in the campaign.
7) In under two years, it was transformed from a state-owned ***l***___ ___ into a successful company.
8) After dealing with all the ***r***___ ___ and finally getting approval we are told that none of the money is forthcoming.
9) Having no particular political ***a***___ ___ ___, he stood for election as an independent candidate.
10) Up to now both sides in the dispute have been cautious, but now ***t***___ ___ ___ ___ and a serious confrontation is expected.
11) Why don't you ***l***___ ___ ___ ___ a bit? Come out with us for the evening.
12) The mayor ***l***___ ___ with her deputy over the plans for the new road.
13) Several Cabinet Ministers have been forced to ***w***___ ___ ___ following the latest Government scandal.
14) Do you still want to do this parachute jump or are you ***g***___ ___ ___ ___?
15) They seem to ***m*** ___ ___ ___every time I meet the conditions which are required.
16) With our credit cards, we've been spending more money than we have. We're going to have to ***b***___ ___ ___ and figure out a way to pay for everything we've charged.
17) Don't suggest that the boss should get rid of one of his secretaries so that we might hire another clerk. The subject of his secretaries is a ***s***___ ___.
18) Amy really has a ***g***___ ___. Everything she plants in her garden grows so well.
19) Constance finally quit her job because the boss asked her to make the coffee and act as a hostess, even though she was hired as an accountant. The ***l***___ ___ came when the boss asked her to go out and buy his family's Christmas presents and then complained because she couldn't get her work done.
20) The ***a***___ ___ for laundry soap is not how well it cleans in hot water, but how well it cleans in cold water.
21) Can you get me a job in your father's company? I know you can do it if you ***p***___ ___ ___.

22) The voters know very little about Mr. Johnson, but he's a ***d***___ ___ and I think he'll win the election.
23) You bought your son a football and he's only six weeks old. Don't you think you're ***j***___ ___ ___ a little?
24) He came home early from work looking worried. He had just gotten a ***p***___ ___, and now he would have to find another job.

Receptive form multiple-choice test:

In this second activity, you will answer multiple-choice questions. Choose only ONE answer. If you do not know the answer, choose the option "I don't know."

1) Unarmed policemen walking the streets late at night are _____.

a) saving face
b) sitting ducks
c) salting away
d) seeing red
e) I don't know

2) He has a reputation as a _____ whose comments sometimes upset Wall Street.

a) lucky dog
b) long shot
c) loose cannon
d) letter perfect
e) I don't know

3) The theatre is a real _____. It costs millions to build and nobody ever goes there.

f) white lie
g) wet blanket
h) white elephant
i) wet whistle
j) I don't know

4) When she failed the entrance exam, she started saying that she never wanted to go to college anyway, but I think that's just _____.

a) sour grapes
b) street smarts
c) sales pitch

d) sitting pretty
e) I don't know

5) The feeling among the _____ of the party is that the leaders aren't radical enough.

a) good eggs
b) gravy train
c) grass roots
d) good Samaritans
e) I don't know

6) This is a total political _____ and an attempt to divert from the main issues in the campaign.

a) red herring
b) real state
c) right hand
d) rock bottom
e) I don't know

7) In under two years, it was transformed from a state-owned _____ into a successful company.

a) lucky stiff
b) last gasp
c) low profile
d) lame duck
e) I don't know

8) After dealing with all the _____ and finally getting approval, we are told that none of the money is forthcoming.

a) red tape
b) running amok
c) real McCoy
d) red-lettered day
e) I don't know

9) Having no particular political _____, he stood for election as an independent candidate.

a) angling for something
b) another think coming
c) arms about it
d) ax to grind
e) I don't know

10) Up to now both sides in the dispute have been cautious, but now the gloves are off and a serious confrontation is expected.

a) the gloves are off
b) they blow their own horn
c) they carry the ball
d) the world is their oyster
e) I don't know

11) Why don't you _____ a bit? Come out with us for the evening.

a) leave holding the bag
b) lie through your teeth
c) lie your way out
d) let your hair down
e) I don't know

12) The mayor _____ with her deputy over the plans for the new road.

a) left behind
b) laid foundations
c) locked horns
d) lost sight
e) I don't know

13) Several Cabinet Ministers have been forced to _____ following the latest Government scandal.

a) walk on air
b) watch the time
c) walk the plank
d) wet their whistle
e) I don't know

14) Do you still want to do this parachute jump or are you _____?

a) going nowhere fast
b) getting cold feet
c) going separate ways
d) getting into it
e) I don't know

15) They seem to _____every time I meet the conditions which are required.

a) move the goalposts
b) make the grade
c) miss the boat
d) make ends meet
e) I don't know

16) With our credit cards, we've been spending more money than we have. We're going to have to _____ and figure out a way to pay for everything we've charged.

a) bend the ear
b) beat the bushes
c) bite the bullet
d) burn the bridges
e) I don't know

17) Don't suggest that the boss should get rid of one of his secretaries so that we might hire another clerk. The subject of his secretaries is a _____.

a) hot seat
b) sacred cow
c) silver screen
d) soft touch
e) I don't know

18) Amy really has a _____. Everything she plants in her garden grows so well.

a) grim reaper
b) good hand
c) green thumb
d) good read
e) I don't know

19) Constance finally quit her job because the boss asked her to make the coffee and act as a hostess, even though she was hired as an accountant. The _____ came when the boss asked her to go out and buy his family's Christmas presents and then complained because she couldn't get her work done.

a) lousy day
b) last straw
c) long face
d) last lap
e) I don't know

20) The _____ for laundry soap is not how well it cleans in hot water, but how well it cleans in cold water.

a) acid test
b) Achilles' heel
c) all hands
d) ancient history
e) I don't know

21) Can you get me a job in your father's company? I know you can do it if you _____.

a) pop the question
b) pull the strings
c) pick his brain
d) pull his leg
e) I don't know

22) The voters know very little about Mr. Johnson, but he's a _____ and I think he'll win the election.

a) dirty trick
b) dead wood
c) dark horse
d) dead ringer
e) I don't know

23) You bought your son a football and he's only six weeks old. Don't you think you're _____ a little?

a) juggling with balls
b) judging the case
c) jogging your memory
d) jumping the gun
e) I don't know

24) He came home early from work looking worried. He had just gotten a _____, and now he would have to find another job.

a) pink slip
b) penny pincher
c) plain sailing
d) pecking order
e) I don't know

Productive meaning recall test:

In this third activity, you will answer questions about the meaning of some idioms. Provide short definitions of the idioms if you know their meanings. If you do not know the meaning, write an "X" in the box.

1) What is the meaning of the expression "sitting ducks"?

_____

2) What is the meaning of the expression "loose cannon"?

_____

3) What is the meaning of the expression "white elephant"?

_____

4) What is the meaning of the expression "sour grapes"?

_____

5) What is the meaning of the expression "grass roots"?

_____

6) What is the meaning of the expression "red herring"?

_____

7) What is the meaning of the expression “lame duck”?

_____

8) What is the meaning of the expression “red tape”?

_____

9) What is the meaning of the expression “an ax to grind”?

_____

10) What is the meaning of the expression “gloves are off”?

_____

11) What is the meaning of the expression “let your hair down”?

_____

12) What is the meaning of the expression “lock horns”?

_____

13) What is the meaning of the expression “walk the plank”?

_____

14) What is the meaning of the expression “get cold feet”?

_____

15) What is the meaning of the expression “move goalposts”?

_____

16) What is the meaning of the expression “bite the bullet”?

_____

17) What is the meaning of the expression “sacred cow”?

_____

18) What is the meaning of the expression “green thumb”?

_____

19) What is the meaning of the expression “last straw”?

_____

20) What is the meaning of the expression “acid test”?

_____

21) What is the meaning of the expression “pull strings”?

_____

22) What is the meaning of the expression “dark horse”?

_____

23) What is the meaning of the expression “jump the gun”?

_____

24) What is the meaning of the expression “pink slip”?

_____

Receptive meaning multiple-choice test:

In this fourth activity, you will answer multiple-choice questions. Choose only ONE answer. If you do not know the answer, choose the option “I don’t know.”

1. Sitting duck

a) a person who is lazy and never takes any exercise
b) a person who likes spending his free time at home, with family
c) a person who is very easy to attack
d) a person who is very passive and likes to be served by others
e) I don’t know

2. Loose cannon

a) a person who is emotionally unstable, who gets upset or excited very easily
b) a person whose behavior is unpredictable, which can have serious or damaging consequences
c) a person who is very energetic and likes to move about a lot
d) a person who behaves in a way that is stupid or not sensible, especially one who is mentally ill
e) I don’t know

3.White elephant

a) something very rare and precious that everyone wants to have
b) something that is useless and expensive to maintain
c) something imaginary, fictitious, hard to believe
d) someone who is socially awkward and does not have many friends
e) I don’t know

4. Sour grapes

a) the attitude of somebody who complains a lot and is always unhappy
b) the attitude of somebody who pretends to dislike something they want because they can’t have it
c) the attitude of somebody who often shows a bad temper and is angry with others
d) the attitude of someone who is no longer young and seems unlikely to marry

e) I don’t know

5. Grass roots

a) the most experienced and important people in a company
b) the least respectable section of society, people for whom one feels contempt
c) ordinary people in society, in contrast with those who make decisions
d) the poorest people in society
e) I don’t know

6. Red herring

a) a subject that people avoid in conversation because it is extremely offensive or embarrassing; a taboo
b) a strange, eccentric person, who likes to shock other people
c) something that takes people’s attention away from the main subject being talked or written about
d) a thing or situation that is difficult or unpleasant to deal with
e) I don’t know

7. Lame duck

a) a thing or situation that is difficult or unpleasant to deal with
b) a person or company that is in trouble and needs help
c) a person who is blamed or punished for the faults of someone else
d) a person who never wins in competitions; a born loser
e) I don’t know

8. Red tape

a) unnecessary official rules and regulations
b) too much talking about unimportant things
c) laws and regulations typical of communist regimes
d) a series of crimes, especially brutal ones
e) I don’t know

9. Ax to grind

a) to have private reasons for being involved in something
b) to have a reason to argue with someone
c) to be in a quarrelsome mood
d) to have business to do with someone
e) I don’t know

10. The gloves are off

a) we are going to have a rest after a day's work
b) we are starting to work seriously
c) we are ready for a fight
d) everything is clear at last
e) I don't know

11. Let your hair down

a) to show sadness and disappointment
b) to reveal a secret
c) to not care about your appearance
d) to relax, enjoy yourself
e) I don't know

12. Lock horns

a) to be on intimate terms with someone
b) to have an affair with someone else's husband
c) to start fighting, arguing with someone
d) to greet someone by hitting their fist with your fist
e) I don't know

13. Walk the plank

a) to be forced to leave your job because something you were responsible for went wrong
b) to leave your job because you have found another, better one
c) to go to the front of a line of people without waiting for your turn
d) to find yourself in a boring, dead-end job which offers you no chance of promotion or development
e) I don't know

14. Get cold feet

a) to be a person who is unfriendly, reserved
b) to be a person who remains calm in stressful situations
c) to start feeling too nervous or afraid to do something
d) to experience the unpleasant state that drug addicts experience when they suddenly stop taking a drug
e) I don't know

15. Move the goalposts

a) do everything you can in order to help someone, achieve something
b) secretly control the actions of other people
c) use your influence to get an advantage for someone
d) to change the rules for something so that the situation becomes more difficult for someone
e) I don't know

16. Bite the bullet

a) to respond to fierce criticism
b) to be hurt by someone else's comments
c) to face a difficult or unpleasant situation
d) to tolerate someone's offensive comments to avoid a heated discussion
e) I don't know

17. Sacred cow

a) something that many people think should not be questioned or criticized
b) something that should be cared for its value
c) something that must not be changed to avoid negative consequences
d) something that was achieved after making a great physical or financial effort
e) I don't know

18. Green thumb

a) the ability to cultivate plants that produce abundant harvests
b) the inefficient care for plants that keeps them small and delay the fruit ripening
c) the ability to take care of plants and make them grow
d) the technique that involves the use of fertilizers to make plants grow faster
e) I don't know

19. Last straw

a) the last step that needs to be taken to accomplish something
b) the action that makes an occasion special
c) the last in a series of bad events that makes it impossible for you to accept a situation any longer
d) the event that causes the end of a personal or professional relationship
e) I don't know

20. Acid test

a) a situation that is difficult to overcome

b) a situation that is difficult, but rewarding since it leads to personal growth
c) a way of deciding whether something is successful or true
d) a way of proving that somebody is not qualified for a task
e) I don't know

21. Pull the strings

a) to use someone's influence to control events or the actions of other people
b) to tell someone that they should wait a moment and not be so excited that they take action without thinking about it first
c) to prevent someone from doing something or make them do something they do not want to do
d) to force someone to tell the whole truth about what happened
e) I don't know

22. Dark horse

a) a person who stands out from the rest after a poor performance in a competition
b) a person who causes trouble in a competition
c) a person taking part in a race or competition who surprises everyone by winning
d) a person whose selfish attitude has a negative impact on a team's performance
e) I don't know

23. Jump the gun

a) to defend yourself from someone who is trying to use violence to hurt you
b) to take action quickly to avoid serious consequences
c) to avoid a danger that is a threat to your physical and emotional well-being
d) to do something too soon, before the right time
e) I don't know

24. Pink slip

a) a letter given to someone to say that they must leave their job
b) a letter to inform that someone has been suspended from work
c) a letter used to inform that someone is going on sabbatical
d) a letter to warn that a job position is at risk due to financial problems
e) I don't know

Background information questionnaire

Gender

- Female

- Male
- Other
- Preferred not to answer

Age group

- 18-20
- 21-23
- 24-26
- 27-29
- over 30

First language (mother tongue)

- Spanish
- Other:

Proficiency level

Have you ever taken an international exam?

- Yes
- No

If you have taken an international exam, which one did you take?

APPENDIX C: Conditions

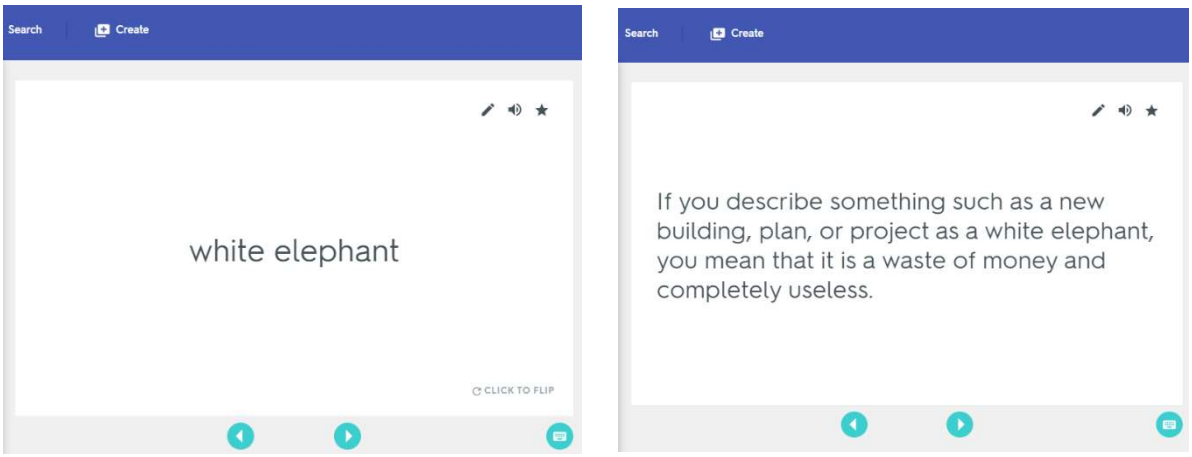


Figure 6. Condition A: definition of idiomatic meaning

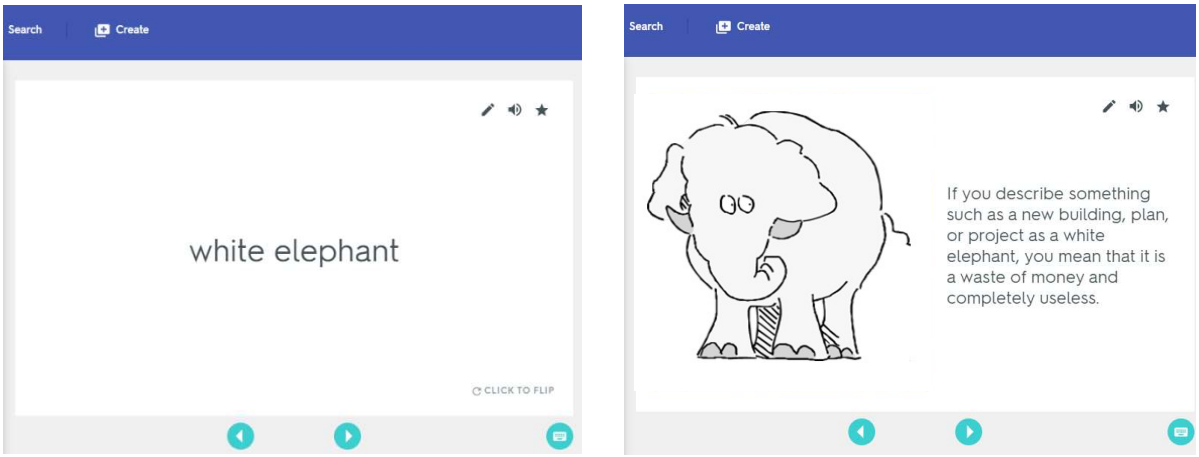


Figure 7. Condition B: definition of idiomatic meaning and picture

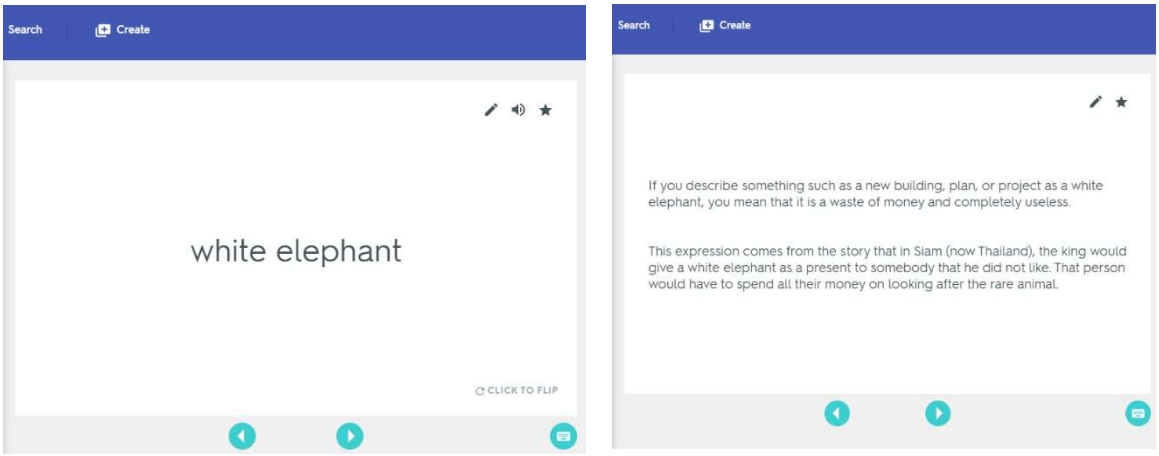


Figure 8. Condition C: definition of idiomatic meaning and etymological note

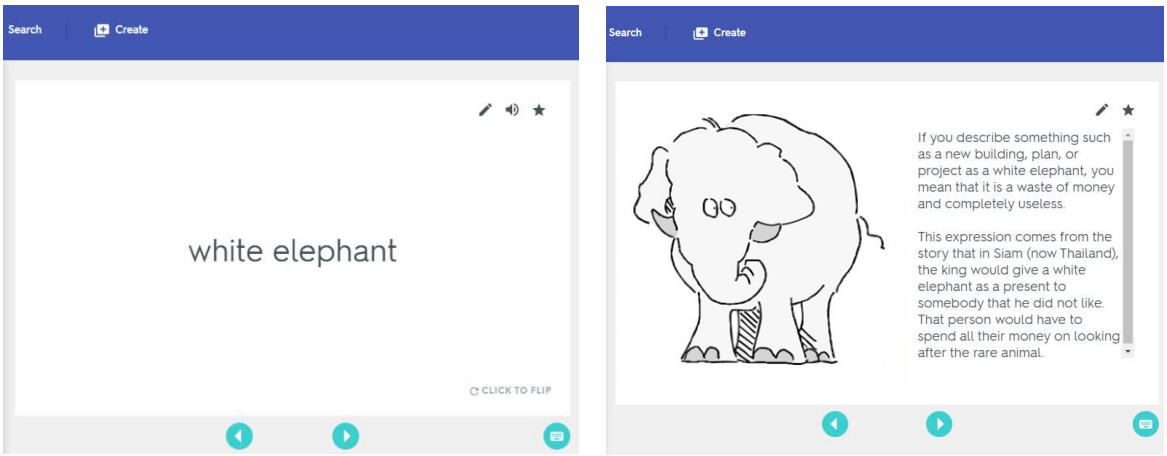


Figure 9. Condition D: definition of idiomatic meaning, picture and etymological note

REFERENCES

REFERENCES

Boers, F., Lindstromberg, S., Littlemore, J., Stengers, H., & Eyckmans, J. (2008).Variables in the mnemonic effectiveness of pictorial elucidation. In Boers, F. and Lindstromberg, S. (Eds.), *Cognitive linguistic approaches to teaching vocabulary and phraseology* (pp. 189–216). Berlin/New York: Mouton de Gruyter.

Boers, F., & Lindstromberg, S. (2009). *Optimizing a Lexical Approach to Instructed Second Language Acquisition*. Basingstoke: Palgrave Macmillan.

Boers, F., Piquer, A-M, Stengers, H., & Eyckmans, J. (2009). Does pictorial elucidation foster recollection of figurative idioms? *Language Teaching Research*, 13, 367–388.

Boers, F. (2011). Cognitive Semantic ways of teaching figurative phrases: An assessment. *Review Of Cognitive Linguistics*, 9(1), 227-261. doi:10.1075/ml.9.1.11boe

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: a framework for memory research, *Journal of Verbal Learning and Verbal Behavior,* 11, 671–684.

Clark, J., & Paivio, A. (1991). Dual Coding Theory and Education. *Educational Psychology Review*, 3(3), 149-210. Retrieved from http://www.jstor.org.proxy1.cl.msu.edu/stable/23359208

Folse, K. S. (2004). *Vocabulary Myths. Applying Second Language Research to Classroom Teaching*. Ann Arbor: The University of Michigan Press.

Gallese, V., & Lakoff, G. (2005) The Brain's concepts: the role of the Sensory-motor system in conceptual knowledge, *Cognitive Neuropsychology*, 22:3-4, 455-479. DOI: 10.1080/02643290442000310

Hulstijn, J. (2001). Intentional and incidental second language vocabulary learning: a reappraisal of elaboration, rehearsal and automaticity in P. Robinson (ed.) *Cognition and Second Language Instruction*, pp. 258-86 (Cambridge: Cambridge University Press).

Irujo, S. (2009). Steering clear: Avoidance in the production of idioms. *IRAL - International Review of Applied Linguistics in Language Teaching*, 31(3), pp. 205-220. Retrieved 21 May. 2018, from doi:10.1515/iral.1993.31.3.205

Lewis, M. (1993). *The Lexical Approach*. Language Teaching Publications.

Lakoff, G. (1987). *Women, fire and dangerous things: what categories reveal about the mind*. Chicago, IL: University of Chicago Press.

Martinez, R., & Schmitt, N. (2012). A phrasal expressions list. *Applied Linguistics*, 33(3), 299-320.

Nation, P. (2013). *Learning Vocabulary in Another Language* (second edition). Cambridge: Cambridge University Press.

Nelson, D.L, Reed, V.S. & Walling, J.R. (1976). Picture superiority effect. *Journal of Experimental Psychology: Human Learning and Memory*, 2, 523–28.

Paivio, A. (1986). *Mental representations.* Oxford: Oxford University Press.

Sadoski, M., & Paivio, A. (2013). *Imagery and text: A dual coding theory of reading and writing.* New York: Routledge.

Pavičić Takač, V. (2008). *Vocabulary Learning Strategies and Foreign Language Acquisition*. Clevedon, UK: Multilingual Matters. Retrieved from https://search-ebscohost-com.proxy1.cl.msu.edu/login.aspx?direct=true&db=e000xna&AN=222245&scope=site

Szczepaniak, R., & Lew, R. (2011) The Role of Imagery in Dictionaries of Idioms, *Applied Linguistics*, 32(3), pp. 323–347, https://doi org.proxy2.cl.msu.edu/10.1093/applin/amr001

Schmitt, N. (2010). *Researching Vocabulary: A Vocabulary Research Manual.* Basingstoke: Palgrave Macmillan.

Wray, A. (2008). *Formulaic Language: Pushing the Boundaries*. Oxford: Oxford University Press.