Electronic Thesis and Dissertation Repository

8-3-2022 1:00 PM

A Study on How Different Glosses Affect L2 Idiom Acquisition

Liting Luo, *The University of Western Ontario*

Supervisor: Dr. Frank Boers, *The University of Western Ontario*
A thesis submitted in partial fulfillment of the requirements for the Master of Arts degree in Education
© Liting Luo 2022

Abstract

Adding glosses to a reading text is expected to be helpful for second (L2) language learners, especially when inferring the meaning of words or expressions is challenging. This project examines the use of glosses to foster comprehension and retention of L2 idioms (e.g., *go against the grain* and *stick to your guns*). More specifically, it compares the benefits of different types of information in glosses: simply clarifying the idiomatic meaning, clarifying the literal meaning from which the idiomatic meaning is derived, or clarifying both. The participants were 37 Chinese ESL learners who read texts with one of the three types of glosses, and then sat immediate and one-week delayed post-tests. Fifteen of them also participated in a stimulated recall interview with the researcher. The post-tests showed no significant difference in the overall effectiveness of the three gloss types for idiom learning. The interview data, however, revealed substantial variation in the ways participants approached the glossed texts, regardless of the reading condition they had been assigned to. The interview data also suggested that the effectiveness of providing information about the literal underpinning of an idiom depends on how easy it is for the individual learner to appreciate the connection between this literal use and the idiomatic meaning.

Keywords: Second language acquisition; idioms; reading with glosses; incidental vocabulary acquisition; mixed-methods research.

Summary for Lay Audience

Second language learners often fail to understand idiomatic expressions (e.g., *follow suit*, *throw in the towel*, *it goes against the grain*, *a wet blanket*) in discourse, including reading texts. One solution is to add glosses to the texts to clarify the meaning of such expressions. Apart from assisting text comprehension, glosses have the potential to help learners remember idiomatic expressions. However, the effectiveness of glossing for the latter purpose may depend on the kind of glosses that are added to the text. An interesting possibility with regards to idioms is to include information not just about the current figurative meaning of the expression but also about the context in which it was originally used literally. For example, the idiom *learn the ropes* (learn how to do a task) has its origin in seafaring, where a novice sailor had to learn how to handle the ropes on a sailing vessel. It is hypothesized that this kind of information can make the expressions more memorable. This study, therefore, compares the effectiveness of glosses with and without notes about the original use of such expressions for learners' comprehension and retention of idioms. Moreover, it is hypothesized that providing glosses *only* with information about the literal use of the expressions will pique the learners' curiosity about the actual meaning of the idioms and engage them in efforts to infer the latter, which may be expected to be beneficial for retention in memory as well.

37 Chinese ESL students from Western University in Canada were divided into three groups. They read two articles that each contained five idiomatic expressions

and answered several content-related questions. The texts were accompanied by marginal glosses about the idioms, but these glosses differed depending on the group the participants were assigned to. One group read the texts with glosses that simply explained the contemporary, figurative meaning of the idioms—corresponding to their use in the texts); another group was given glosses that only explained the original, literal use of the expressions; and the third group was given glosses which presented both types of information. They were asked to complete a number of tests shortly after the reading activity and again one week later to compare how much idiom learning happened under the three reading conditions. After the final tests, 15 participants (5 from each group) were invited for an interview in which they were asked to recall how they approached the reading texts, the glosses, and the idioms. The test results showed no relationship between reading conditions and learning gains. However, the interviews revealed substantial variation in how the participants approached the reading tasks and the information about the idioms, often regardless of the type of glosses they were given. The nature of the idioms also helped to explain why a certain type of gloss (e.g., literal-origin only) was more suitable for some items than for others. The study thus illustrates the usefulness of using different sources of information (in this case test and interview data).

Acknowledgments

I spent 16 months on this thesis and now it is finally done. On this page, I would like to mention a few people and say, "Thank you!" to them. I am a very emotional person. When writing this, I always unconsciously shed tears. So, please listen to me.

First, I would like to express my sincere appreciation to my GREATEST supervisor in the world, Dr. Frank Boers! Thank you for taking me into the world of academia and giving me countless inspiring ideas and insightful feedback. Thank you for not only caring about my academic achievements but also about my life. I know sometimes I am quite "annoying", but the smile in your eyes always comforts me.

Then, I would also like to thank the 37 participants, you are not only my colleagues but also my friends. Without your selfless and voluntary support, this study could not have been completed.

I would like to give my special thanks to Dianyuan (Eden) Yi. Thanks for piloting this experiment with me and giving me valuable feedback and suggestions. Also, thanks to Xiaohuge's food, which helped me not to feel too homesick when I set foot on this land again.

I thank my colleagues Xi Yu, Yilin (Linda) Wang, Mengxue (Alyssa) Li, and Shanshan (Elena) He. Thank you for providing me with technical support! Special thanks go to my special little friend, Bubble! Thank you for comforting me during my saddest nights.

Finally, I would like to thank my parents and grandparents. Thank you for your unreserved dedication and understanding. Thank you for letting me know, no matter how far I go. For as long as I look back, I can see you. Thank you, my boyfriend, Yanchang, for embracing all my emotions and protecting me. You are always my most determined choice (I know you can read English, so I will not translate this part for you. Hope you still can enjoy it). [最后，我想感谢我的父母和外公外婆还有爷爷。 谢谢你们对我毫无保留的奉献和支持，谢谢你们让我知道，无论我走的有多远，只要回头，你们一定在。]

Table of Contents

List of Tables

List of Figures

List of Appendices

Chapter 1 Introduction

Reading second language materials has been considered an indispensable way to increase vocabulary knowledge (e.g., Krashen, 1989). However, successful word learning cannot be guaranteed under every natural context (Laufer, 1997). Thus, some reference methods such as glossing have been suggested to increase the possibility of learning from reading. Glosses (also called annotations) are added to a text and used as clarifications to support comprehension (Boers, 2022). In the case of reading print text, glosses can be located at three places: in the margin of the text, in the text next to the target item, and in the form of a glossary at the end of the text. Even though originally glossing was intended to help students understand texts autonomously (Davis, 1989; Holley & King, 1971; Jacobs, 1994; Lomicka, 1998), exploring the potential of glosses for lexical learning has become the principal focus in more recent research.

A great number of studies have shown a positive effect of glossing on lexical acquisition (Cheng & Good, 2009; Poole, 2012; see Yanagisawa et al., 2020, for a meta-analytic review). In the meantime, many researchers (e.g., Barfield & Gyllstad, 2009; Boers & Lindstromberg, 2009; Siyanova-Chanturia & Pellicer-Sanchez, 2018; Wood, 2010) have started examining ways of helping L2 learners to acquire not only single words but also multiword items (henceforth MWI). However, apart from a few exceptions such as Peters (2012), hardly any studies interested in the acquisition of multiword items from reading have focused on the use of glosses. Glosses can be

expected to be helpful for second (L2) language readers especially when inferring the

meaning of the expression is challenging, as is the case with idioms. Idioms are

figurative expressions whose meaning does not follow straightforwardly from the

basic meanings of their constituent parts. For example, *a wet blanket* originally

referred to something that can be used to extinguish a fire. Its contemporary idiomatic

meaning is figurative and refers to someone who spoils other people's excitement.

Another example is *a shot in the arm*, which referred originally to an injection of

medicine or drugs to make you feel better. The current figurative meaning is

something that helps you or energizes you.

This study will focus on the potential benefits of using glosses for idiom

learning. More specifically, the purpose is to investigate which type of glosses (see

further below) works better for the acquisition and retention of L2 idioms. Moreover,

the findings of this research may raise teachers' awareness of the challenges posed by

idioms and help them in their future teaching.

Chapter 2 Literature Review

In this part, the reasons for focusing on idioms as a subclass of MWIs will first be clarified. Next, the benefits of reading with glosses for vocabulary acquisition will be discussed. Then, I will turn to the rationale for comparing the benefits of different types of glosses about idioms—glosses that present readers with the literal underpinning of the expressions versus glosses that present readers only with the abstract or figurative meaning of the expressions as used in the text.

2.1 What are Multiword Items?

During the last two decades, applied linguistics have developed a strong interest multiword items (e.g., Conklin & Schmitt, 2012; Foster, 2001). Research foci vary, and there will be some variation in the labels and definitions of these items. Wray (2002) used "formulaic sequence" as a label and pointed out two key features: first, a formulaic sequence should include more than one word; second, it appears to be stored in and recalled from memory as a whole "prefabricated" unit, instead of being created word by word with the help of grammar.

In this research proposal, MWIs is used as an umbrella term to cover a wide range of above-one-word expressions. Here are some examples of multiword items: (1) lexical bundles (e.g., *all in all*), (2) lexicalized sentence stems (e.g., *it cannot be denied that…*), (3) collocations (e.g., *lose weight*), (4) compounds (e.g., *contact lens*; *bank account*, (5) phrasal verbs (e.g., *drive away; give up*), (6) proverbs and

aphorisms (e.g., *When the cat's away, the mice will play; all good things come to an end; practice makes perfect*), (7) binomials (e.g., *pros and cons*; *bread and butter*), (8) standardized similes (e.g., *cold as ice; good as gold*) (9) idioms (e.g., *go against the grain*; *follow suit*), and (10) conversational formulae (e.g., *How are you doing?*).

2.2 Why are MWIs Important?

MWIs serve different purposes in language and language learning. First, mastery of MWIs boosts both receptive and productive fluency (Boers, 2020). For the receptive aspect, with the knowledge of MWIs, speakers will feel more confident to make predictions in discourse (Boers & Lindstromberg, 2012). For instance, when advanced language users hear *so on and so*, they will expect *forth* to complete the sequence. Similarly, learners familiar with discourse about language education will be able to anticipate *Languages* after hearing or reading *Teaching English to Speakers of Other*. For the productive aspect, if learners can retrieve MWIs from memory successfully, they will directly produce a multi-word chunk "holistically", as a single lexical item with its own distinct meaning of function (Thomson et al., 2017). For example, learners may use *all in all* or *as a result* to conclude something; *on the one hand…on the other hand* will be used to compare and contrast things.

Second, MWI knowledge benefits learners' accuracy. Cobb (2018) has suggested that students who lack MWI knowledge are more likely to produce non-standard or "wrong" expressions. To some extent, learners' first language influences second

language acquisition, which is known as L1 transfer (or L1 interference). For example, Chinese L2 learners find it hard to choose between *do* and *make* as these two verbs have a single counterpart in Chinese. Thus, instead of using the correct collocation "to make a decision" they may produce "to do a decision". This can be problematic, because some studies have revealed that collocational accuracy is a strong predictor of L2 proficiency ratings (e.g., Crossley et al., 2015)

Third, Boers et al. (2006) demonstrated that learners who master a wide range of MWIs produce discourse that is perceived to be lexically rich, approximating the discourse of L1 users. Use of diverse MWIs, especially ones that include relatively low-frequency words, is strongly associated with perceived proficiency (e.g., Bestgen, 2017). Learners' use of MWIs with a metaphorical meaning (such as figurative idioms) has been shown to be particularly strongly associated with proficiency ratings (e.g., Hoang & Boers, 2018).

2.3 The Challenges of Learning MWIs

Many researchers have recognized that MWIs are hard for L2 language learners to acquire, especially for those who do not enjoy much exposure to the target language, such as learners who live in an EFL context (e.g., Granger, 1998; Laufer & Waldman, 2011; Li & Schmitt, 2010; Nesselhauf, 2003; Siyanova & Schmitt, 2008). Moreover, some researchers have pointed out that even high-level language learners can only use a limited number of MWIs and there is no guarantee that every use is

correct (Altenberg & Granger, 2001; Laufer & Waldman, 2011; Nesselhauf, 2005). There are several explanations for this.

First, several authors (e.g., Boers, 2020; Martinez & Murphy, 2011) have noted that many MWIs are composed of highly frequent and thus highly familiar words. Compared to less familiar (more "difficult") words, these words often fail to attract attention, and attention is known to play a crucial role in language learning (Godfroid et al., 2013; Schmidt, 2001; Tulving & Kroll, 1995).

Second, although MWIs *as a class* are ubiquitous in natural discourse, it is rather unusual to meet the same individual MWIs frequently in one piece of text (with the exception of a few lexical bundles and "fillers", such as *you know* and *kind of*). Webb (2007) suggested that if learners meet the same lexical items many times, they are more likely to be acquired. However, the same MWIs seldom show up repeatedly in the same conversation or text. Boers and Lindstromberg (2009) illustrated this by screening a novel for recurring verb-noun collocations. They found that, for example, *tell the truth* appeared only once in 100 pages of the novel. The novel happened to be a police story, and this type of story could be expected to include more instances of this collocation. Beyond a few high-frequency MWIs (e.g., Shin & Nation, 2007) students who do not have regular exposure to rich L2 input may not have many opportunities to encounter the same MWIs in a short time. This is especially hard for learners in EFL contexts. Beyond their textbooks and classroom activities, EFL learners have little opportunity to meet the same MWIs repeatedly in L2 discourse.

Third, it is difficult for learners to recognize MWIs in discourse without instruction. With the help of empty spaces before and after a word in written texts, learners can distinguish each word; however, identifying MWIs is not as easy as identifying single words in a text (Eyckmans et al., 2007; Lindstromberg et al., 2016) because it is not clear what word sequences constitute a discrete semantic or lexical unit. Moreover, even if they feel that a certain word sequence must be an expression, L2 learners may not be able to work out what it means. For example, in the song *Lala Land*, there is a line in the lyrics "I feel a little *under the weather*". Learners may understand the meaning of the individual words, but the meaning of the expression as a whole may remain elusive. Besides, in the case of songs, it is difficult for learners to decide whether a given phrasing is conventional or the product of creativity (akin to poetry). It stands to reason that one needs to meet the same phrase several times in various contexts to develop an intuition that it is a conventional phrase, and thus worth remembering.

Fourth, it seems that teachers and learners are more inclined to teach and learn single words than MWIs (Bui et al., 2020; Peters, 2012). In Chinese English textbooks, for example, the number of single words in the final word list is much larger than the number of MWIs. This lack of appreciation on the part of teachers and textbook authors of the importance of mastering MWIs may be another reason for the slow acquisition of such items by L2 learners.

The fifth challenge in learning MWIs is one I already alluded to previously: Their meaning can be quite elusive, even if they are made up of familiar words (e.g., Boers & Webb, 2015; Littlemore et al., 2011; Martinez & Murphy, 2011). This holds true especially for a particular sub-class of MWIs, notably idioms. For instance, the idiom *jump the gun* ('act too soon') is hard to infer the meaning of, especially if a learner associates *gun* with a weapon instead of its less common reference to a starting pistol that is used in racing contests (which is where this idiom comes from). A learner may stand a slightly better chance of guessing the meaning of *toe the line* ('follow the rules'), because the combination of *toe* and *line* may conjure up the image of track athletes lining up at the starting line. Still, there is no guarantee of this, because the word *line* is polysemous and could refer to various things (a line of text, a fishing line, etc.) the image of which will not help to interpret the expression. Often, then, learners' unassisted interpretation of idioms will depend on the availability of supporting context (e.g., Cooper, 1999). Unfortunately, research on vocabulary acquisition from reading has demonstrated that clear contextual clues are often missing in authentic discourse (e.g., Nassaji, 2003), and that learners do not always make good use of such clues on the rather rare occasions that they are available (Bensoussan & Laufer, 1984).

2.4 Why Focus on Idioms?

Although second language learners may understand the individual words that make up an idiom, they will often fail to understand the meaning of the expression as a whole because this does not follow directly from adding up the meanings of the constituent words. There is evidence that learners tend to activate literal readings of the constituent words when they process *known* idioms (Cieślicka, 2006, 2012), but this will not help interpretation of *new* idioms if these are literal readings that are not actually what the idioms were derived from. Oftentimes the first literal meaning likely to come to a learner's mind will not be a helpful clue. This is the case, for example, for phrases such as *against the grain*, *to follow suit*, and *to go out on a limb*, where learners are likely to think of grain in the context of farming, a suit as clothing, and a limb as a body part. However, the references are to the grain of a piece of wood (i.e., the direction of its fibres), a suit of cards (e.g., diamonds) in a card game, and a limb of a tree (i.e., a branch).

One might argue that idioms are a marginal phenomenon in language and thus not deserving of much attention in teaching and learning. As a class, however, idioms occur quite frequently in L1 users' everyday discourse (Boers & Lindstromberg, 2009). Based on idiom counts in the *Word Banks* corpus (i.e., the corpus used for the *Collins Cobuild* dictionaries), Boers and Webb (2015) estimate that, on average, 2,400 idioms occur per million words of English discourse. It is worth specifying that this count concerns just expressions listed in an idiom dictionary, while there are other

(semi-)fixed phrases (e.g., phrasal verbs) whose meaning can elude L2 learners. In

sum, idioms are challenging for non-native speakers because they are bound to meet

them, and when they meet them they will probably do not understand them. For

example, Martinez and Murphy (2011) had low-intermediate EFL learners read texts

that consisted exclusively of high-frequency words likely to be familiar to these

learners, but with a number of these words figuring in idiomatic expressions, such as

he's over the hill ('he's beyond his prime'). Despite available context to the contrary,

many of the learners interpreted this as "he lives on the other side of the hill".

Littlemore et al. (2011) examined the extent to which international students at a

university in the UK misunderstood their lectures and found that most of their

misunderstandings concerned the lecturers' use of metaphorical language, including

figurative idioms. Clearly, figurative language (including idioms) constitutes an

obstacle to discourse comprehension, and so there is a need to help L2 learners to

overcome this obstacle. A straightforward approach may be to regularly clarify the

meaning of idioms as they are encountered in texts, and in doing so help learners

build receptive knowledge (i.e., comprehension) of idioms they are likely to meet

again. Clarifying the meaning of idioms can be done through glossing.

2.5 What are Glosses?

   Nation (2001, pp. 174-175) defines gloss as "a brief definition or synonym of

unknown words provided in the text in the L1 or L2". Glosses may appear in the

margin area next to a text, at the bottom of a page or at the end of a text as a glossary.

If learners read on a computer, glosses can also be inserted as hyperlinks and readers

can easily access them with a click of the mouse.

2.6 Why Read with Glosses?

Glosses can support reading by (a) introducing the meaning of target words to

learners during bottom-up processing (Gettys et al., 2001), (b) preventing students

from making incorrect inferences (Paribakht & Wesche, 1999), and (c) reducing the

need for dictionary lookups by giving readers easier access to word meanings.

According to the literature on L2 acquisition, lack of "noticing" is one of the major

explanations for unsuccessful learning. Schmidt's (1995) *Noticing Hypothesis*

suggested that conscious attention is necessary for learning, and noticing is generally

the first stage of learning. Sometimes, in the process of reading, the reader may not

notice new words, and so vocabulary learning will not happen. Some researchers

(e.g., Yoshii, 2006; Nation, 2001) suggested that glossing could be considered as one

of the handiest tools to help readers notice (and understand) unfamiliar words.

Apart from prompting readers to take notice of the few words in a text they do not

know yet (or to take notice of *new uses* of already somewhat familiar words), glosses

may also be indispensable for text comprehension, in particular if the text contains a

large number of unfamiliar words (e.g., Paribakht & Wesche, 1999; Parry, 1997;

Watanabe, 1997). Moreover, compared to checking a dictionary while reading, using

a gloss is easy, time saving and minimizes the interruption of the reading flow (Ko,

2005; Nation, 2001). Thus, glosses are a convenient means for helping learners read

L2 texts autonomously.

2.7 Glosses for Vocabulary Acquisition

The original principal purpose of glossing was to support text comprehension.

Over time, however, researchers have shifted their interest to how well glosses

support vocabulary learning (e.g., Watanabe, 1997). In addition, a substantial number

of studies have compared the effectiveness for vocabulary learning of different types

of glosses. For example, Jacobs et al. (1994) examined L1 glosses and L2 glosses.

The best results were obtained with L1 glosses, although the results of subsequent

studies suggest that L2 glosses can be as effective provided the learners are at a

relatively high proficiency level and find the L2 glosses easy to understand (Boers,

2022, for a critical review). In any case, although L1 glosses may be easier for

learners to take in, there are of course many situations where glosses need to

accommodate readers with diverse L1s, and so brief but transparent definitions in the

target language will often be required.

Another comparison of glosses concerns single textual glosses, single picture

glosses and text plus picture glosses (e.g., Kost et al., 1999; Yoshii & Flaitz, 2002).

The results suggest that students tend to benefit more from text plus picture glosses

(even if only because pictures attract attention whereas words are often overlooked),

although there are a few exceptions (e.g., Acha, 2009; Boers et al., 2017). A practical

problem with pictorial glosses is that they can take up a lot of space (in the case of

printed reading materials), which means that they need to be used economically.

Yet another comparison regards multiple-choice glosses and single-meaning

glosses. Researchers have been concerned that standard glosses may not leave a

profound memory as learners "passively" take in the information and then continue

reading. Therefore, Hulstijn (1992) proposed using multiple-choice glosses, which

include interpretation challenges and thus invite more cognitive engagement with the

glosses. A number of studies (e.g., Nagata, 1999; Yoshii, 2013) have explored this

possibility, but the results have been rather mixed, owing to the interplay of many

variables (e.g., text length, density of unknown words, the readers' proficiency level,

and the reading purpose), and—importantly—the risk of learners choosing a wrong

meaning from among the options presented in the multiple-choice gloss.

Regardless of type of gloss, there is now a large body of evidence that reading

glossed text leads to better word learning than reading the same texts without glosses

(Boers, 2022; Yanagisawa et al., 2020, for reviews), at least according to tests

administered shortly after the reading activity (evidence from delayed tests is weaker

due to attrition over time, e.g., Zhang & Webb, 2019). However, there are very few

studies on glossing that focus on MWIs, not to mention idioms. As illustrated

previously, idioms pose comprehension problems and glossing appears a

straightforward means of overcoming those problems, while at the same time

fostering receptive knowledge of the expressions that will be helpful should the expressions be met again. Thus, it is worth exploring how glosses affect idiom learning. Moreover, as will be explained next, glosses for idioms can be designed in ways that promote cognitive engagement as well, but unlike multiple-choice glosses without the need to present learners with wrong meanings.

2.8 Connecting Literal and Abstract Meanings

Many lexical items are polysemous. Polysemy means that a word or expression has more than one meaning. Usually, a polysemous item has a root literal meaning as well as more abstract, figurative meanings derived from it. For example, *hand* literally means a part of the human body beyond the wrist. However, when people say, "give me a *hand*", it means "help" which is an abstract meaning. In "He's a new *hand*", it means performer or worker. It is not unusual for second language readers to encounter the abstract or figurative use of a word without being aware of its original, literal underpinning. For example, students specializing in the area of economics may meet the expression "economic recovery" without being aware that recovery is also used in the context of recovery from an illness. Might it benefit learners' retention of abstract lexical items encountered in texts if the accompanying glosses pointed to their literal underpinnings?

In Boers (2000), a group of students who majored in commerce and economics read a text which included several polysemous items. One group of students received

a glossary of these words explaining the abstract meanings of the items as used in the text. For example, *bail out* refers to giving financial aid; *shift tack* refers to changing policy. For the experimental group, participants had a glossary with the words' original, literal meanings: *bail out* means keeping a sinking boat afloat by throwing out the water that has come in; *shift tack* means turning a sailing boat around so that the wind catches the other side of the sail. Three days later, the students were given a new text with blanks for them to complete with the words or phrases they had in the previously annotated text. The students who only received the literal meanings in those annotations obtained better scores in this test, even though the lexical items were used in their abstract meanings again in the new text. Boers (2000) speculated that the students were possibly puzzled by the literal-meaning annotations and needed to invest effort in establishing the connection between the literal meaning and the abstract, figurative use of these words and phrases in the actual reading text. It is generally believed that cognitive effort, such as inferencing, invested in a learning task benefits retention (e.g., Mondria, 2003). In addition, awareness of the literal underpinnings probably made these lexical items more "imageable", which likely made the items more memorable (e.g., Paivio, 1986).

The experiment by Boers (2000) had a number of shortcomings, however. One is that the students were not asked if they perhaps already knew some of the lexical items, and so it is not entirely certain if the two participant groups were equivalent in terms of prior knowledge. Another is that it is theoretically possible to complete

blanks with previously seen words even though one is not entirely sure of their meaning. In addition, the above explanations for the better recollection of figuratively used words and phrases after reading their literal meanings must remain speculative. After all, the students were not asked how they had dealt with the information in the glossary, or even if they had read these annotations in the first place.

The study proposed here could be considered a conceptual replication of Boers (2000), but with a tighter focus on idioms and an expanded research design. For example, while Boers (2000) compared the effectiveness of literal-meaning and abstract-meaning glosses, that study did not examine the effect of glosses which present both the literal and the abstract meanings. While such combined information no longer entails the interpretation effort invited by literal-meaning-only glosses, it ensures correct encoding of the abstract, idiomatic meaning, while still stimulating mental imagery.

Chapter 3 Research Questions

To reiterate, there is a large body of evidence that learners benefit more from reading with glosses than from reading without glosses (Boers, 2022; Yanagisawa et al., 2020, for reviews). It seems reasonable to expect this benefit to extend to glosses about idioms, since an idiom functions as lexical unit, akin to a single word. What is far less clear, however, is whether the nature of the information provided in the glosses makes a difference to learners' retention of the meaning of such phrases. This study therefore seeks answers to the following research question:

Is there a difference in the effectiveness of three types of glosses, that is, (a) presenting only the literal meaning that underpins the contemporary figurative meaning, (b) presenting only the figurative (or abstract) meaning, and (c) presenting both meanings, for L2 readers' retention of figurative idioms?

Based on the results in Boers (2000), I expect reading with literal-meaning glosses to be more effective for learners' retention of the idioms than reading with abstract meaning-glosses alone, owing to the effort invited to infer the actual meaning of the idiom aided by the hint about its literal underpinning. However, it is hard to make a prediction when it comes to the effectiveness of the combination of abstract plus literal meaning glosses since this condition has not been put to the test before in research on glossed reading. While the enhanced imageability of the idioms thanks to information about its literal underpinning is expected to be beneficial for retention, presenting it together with the actual abstract, idiomatic meaning does not induce as

much cognitive effort as in the case of glosses where only the literal meaning is given.

So, while the combined glosses may be expected to be more effective for the retention

of the idioms in memory than glosses that only clarify the abstract, idiomatic

meaning, it is less easy to predict whether they will also be more beneficial than

glosses that only provide the literal underpinnings of the idioms. Even when it comes

to the comparison with abstract-only glosses, it is not guaranteed that the elaborate

glosses in which both abstract meanings and literal underpinnings are elucidated will

have the greater effect, because learners may prefer the briefer, to-the-point glosses,

and feel reluctant to interrupt the flow of reading by studying overly lengthy glosses.

Chapter 4 Methodology

4.1 General Research Design

This is a mixed-methods study, using a classroom experiment with a between-participants research design, followed by interviews. The whole study was spread over four weeks and comprised a reading task supported by different types of glosses, an immediate post-test, a delayed post-test, and one-on-one interviews. Figure 1 presents the components of each stage and their relationships.

There are some possibilities to compare the effectiveness of different glosses by using a within-participants design. For example, one could have participants read 3 short texts, each text with a different type of glosses. However, we did not choose this based in the following two reasons: (a) seeing different kinds of glosses would have raised awareness of the purpose of the experiment and (b) it would have been difficult to control for features of the target idioms and their contexts that influence comprehension. Even though one could opt for a counter-balanced design (where some students red text A with literal glosses and text B with abstract glosses, while others read text A with abstract glosses and text B with literal glosses, etc.), it would have been hard to implement this in this current study since it investigates 3 reading conditions, and so a counterbalanced within-participant approach would have made the design overly complex, possibly requiring more participants for this to be feasible.

Participants' scores on the post-tests serve as the quantitative data. The learning outcomes as gauged by the post-tests were the dependent variable, and the

independent variable was the different reading conditions: reading with literal-

meaning glosses, reading with abstract-meaning glosses, and reading with both

abstract-meaning and literal-meaning glosses. The participants' responses in the

interview are qualitative data which helped me to interpret the quantitative data. The

main purpose of the interviews was to explore if and how the participants engaged

with the glosses. Idioms not only have abstract meanings but also have literal

meanings and abstract usage is derived from the original, literal meanings through

metaphorization. Thus, students who read the texts with glosses which only present

the literal meaning of the idioms would need to invest some effort in working out the

connection between the literal meaning and the contextual, figurative use of the items

(Boers, 2000). Through the interview I wished to ascertain if the students managed to

establish the connections between the literal and the figurative meanings.

Figure 1

Different Stages of the Research Design
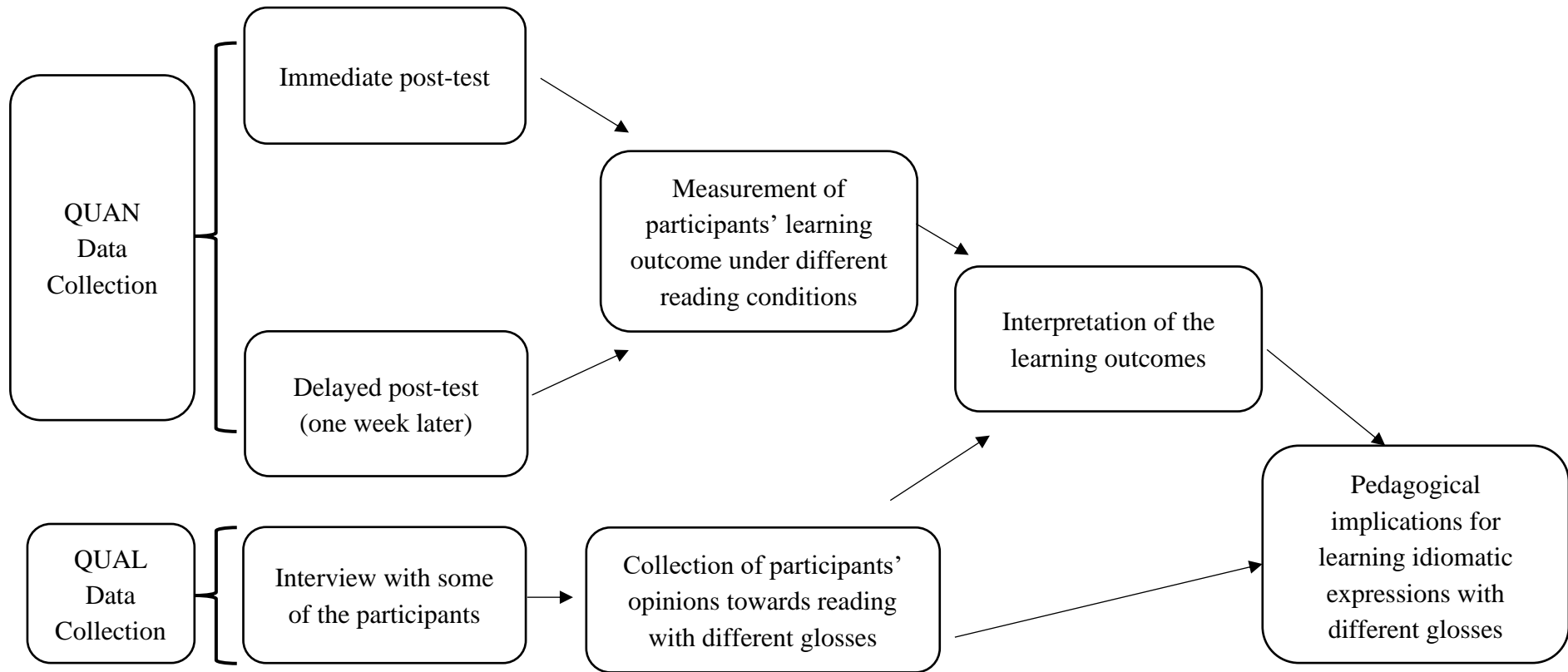
4.2 Participants

There was a total number of 37 participants recruited in this study, 2 male and 35 female students. They were all international students from China studying at *Western University* in Canada. Most of them (n = 26) were enrolled in the MPEd TESOL program. Because this was a rather small number, additional participants were recruited from related programs. The additional participants were MA students (n = 9) and PhD students (n = 1) in the field of Applied Linguistics. The latter were all previous graduates from the MPEd TESOL program. They were all immersed in an English-speaking environment and thus likely to benefit from expanding their knowledge of idiomatic expressions. All participants had been accepted in their respective programs in the Faculty of Education at Western University, and so they met the minimum language requirement—6.5/9 overall score with no individual score below 6 on IELTS; and for TOEFL a minimum of 550 for the paper-and-pen version, and 213 for the computer version, equivalent to B2 (upper-intermediate) of the *Common European Framework of Reference* (CEFR) scale. However, the participants' language proficiency was likely higher as they had been immersed in an English-speaking environment since applying for entry into their respective programs.

4.3 Materials and Instruments

4.3.1 Reading Texts

Two short texts (see Appendix A-F) on different topics were used and participants in all the three conditions received the same reading texts (approximately

600 words each). Participants came from different academic backgrounds prior to enrolment in the TESOL and Applied Linguistics programs and may therefore have different fields of interest and background knowledge. Interest in and familiarity with the topic of a text are known to influence text comprehension and vocabulary uptake (e.g., Pulido, 2004). By using texts on different topics, the probability of this affecting the overall results was reduced. Neither text was highly technical, however. The two articles were slightly modified to keep the essays free of too many difficult words (except for the target idioms).

Text 1—*What Is the Nag Factor?* (https://www.psychologytoday.com/us/blog/the-age-overindulgence/202110/what-is-the-nag-factor) was taken from *Psychology Today.* According to Cobb's (n.d.) Lexical Tutor software (https://www.lextutor.ca/vp/comp/), 94.9% of the words in this article belong to the 3,000 most frequent word families of English, and 98.1% coverage is reached when words belonging to bands up to K-6 are included.

Text 2—*The Earth Is on Fire* (https://www.scientificamerican.com/article/the-earth-is-on-fire/) was selected from *Scientific American*. 95.2% of the lexical item in the text belong to K-1 to K-3, and 98.5% coverage is reached with words from bands up to K-6.

According to these lexical profiles, the reading difficulty of the two articles was similar, and they should be readable for the participants, as, at their level of proficiency, they were almost certainly able to understand at least 95% of the running

words. Besides, some of the words belonging to lower-frequency bands (Table 1) were clarified in glosses.

To make sure that the participants would be reading the text for its content, they were told to answer content-related questions. Similarly, in order to avoid participants becoming aware that the study was specifically about idioms, I annotated some of the lower-frequency single-word items in the texts as well (see Table 1).

Table 1

Lower-frequency Words in Reading Materials Clarified in Glosses

Difficult words	Material	Freq. Level	Gloss in the text
Nag	One	K-6	keep asking someone to do something
Vending (machine)	One	K-10	you can buy drinks, snacks by putting coins into it
Resilience	Two	K-6	the ability to recover after difficulties
Solace	Two	K-9	comfort, hope, consolation

4.3.2 *Answer Sheets*

Participant groups received three answer sheets for the immediate post-test and two answer sheets for the delayed post-test. Each answer sheet was collected when the students indicated they had finished before the next answer sheet was handed out. The answer sheet with content-related questions was given together with the texts and required the participants to express how much they agreed with given statements on a six-point scale questions (see Figure 4). Even if students did not understand the

meaning of the idioms, they should still be able to answer these questions. In other

words, these questions were not directing the readers' attention specifically to the

idioms. After finishing reading and answering these questions, this answer sheet and

the text were collected. Then, a gap-fill test was handed out to the participants. This

presented the participants with excerpts from the text but with blanks replacing key

words of the idioms for the participants to complete. Finally, a test with meaning-

explanation questions was given to students after collecting answer sheet 2. This test

required the participants to recall the meaning of the idioms presented in a sentential

context from the original text. One week later, the students were given a second

meaning-recall test (but without contextual support), followed by a meaning-

recognition test (a multiple-choice test asking participants to match the idioms to their

correct meanings).

Figure 2

Example of Content-related Question

Content-related Questions

In this article, five basic conditional principles are provided to parents. Please read it
again and select the response that best indicates your opinion about the statements
below.

1. *Almost everything is learned.*

	1	2	3	4	5	6	
Totally Disagree	○	○	○	○	○	○	Totally Agree

4.3.3 Rationale for the Post-test Design

This experiment was primarily focused on students' incidental learning outcome (especially meaning recall and recognition) of target idioms under different reading conditions and the retention of knowledge over time. Participants received some content-related questions together with the reading tasks, and they could answer those questions while reading the text. Immediately after the reading tasks, the participants first took a post-test requiring them to reproduce the keywords of idioms in the original context with blanks left to fill in the words (see Figure 3). Next, they were asked to explain the idioms' meaning. (The tests also included the low-frequency words that had been glossed as well, but these were excluded from the analyses, because those glosses were identical across the three conditions.)

In the delayed post-test, the participants were first given a meaning recall test (no longer contextualized) and a meaning recognition, multiple-choice test (see Figure 6). The meaning recognition test was added to the delayed post-test in case of a floor effect in the scores on the meaning recall test. A substantial amount of attrition is likely to happen in the interval between an intervention and the delayed post-test. Since meaning-recall tasks are quite challenging, adding a meaning-recognition task (an easier kind of test) is a way of capturing between-group differences (if any) where results of the more difficult recall test (meaning-recall test) might be too low for a difference to be noticeable. In the delayed post-test, the students needed to recall or choose the most suitable answer without the help of the discourse context in which they encountered the idioms the previous week. Omitting the original text avoided the

possibility that students guess the meanings from the context rather than retrieving

them from memory.

Figure 3

Example of Questions in the Immediate Post-test

Gap-fills

Complete the sentence by filling in the blanks. Each blank stands for one word.

Children use pester power to persuade parents to buy them things that they don't really need. The nag factor not only gets parents to buy things but is also used to manipulate parents to do things they normally would not. A 2002 survey found that most children repeatedly request an item until their parent **throws in the** _____. Some will ask more than 50 times for a particular product.

Meaning Explanations

Please read the sentences and **explain the meaning of the word or phrases in bold**.

2. A 2002 survey found that most children repeatedly request an item until their parent **throws in the towel**.

   _____

   _____

   Did you know this phrase before?

   ☐ Yes

   ☐ No

Figure 4

Example of Questions in the Delayed Post-test

Meaning Explanations

Please write down the meaning of the following word/phrases.

1. Throw in the towel

_____

_____

Multiple-choice Test

Please choose the most suitable meaning for the following words/phrases.

1. Throw in the towel
   - ☐ Refuse someone's request
   - ☐ Give up doing something
   - ☐ Postpone something to deal with something else
   - ☐ Make someone angry

Note that the meaning-recall test ("meaning explanations") was collected before the meaning-recognition test ("multiple-choice test") was handed out. The tests about the idioms were not announced beforehand, thus situating the study within the realm of "incidental" vocabulary acquisition (e.g., Webb, 2020). To check for prior knowledge of the idioms, the learners were asked to state whether they already knew any of the idioms before the experiment (see Figure 3). Data regarding items that were already known were excluded from the analyses. As the principal focus was on the learning gains concerning the idioms, performance on the content-related questions was not considered in this study because these questions were only created

to make sure participants were reading the text for its content. Besides, these

questions elicited opinions rather than assessing text comprehension.

4.3.4 Interview Protocols

I randomly invited 15 students (five students from each treatment group) for a

stimulated recall interview, to talk about how they recalled/recognized the meanings

(i.e., to inquire about their episodic memories of the reading text and its glosses).

Questions were slightly different for the different condition groups (see Appendices

M, N, and O). However, there were 9 questions in total, one of which was a closed

question, and 8 were open-ended questions. All interviews were conducted online and

recorded via Zoom. The interview did not exceed 20 minutes for any student.

Whether the participants were able to make a link between the idioms' literal meaning

and their figurative meaning was of particular interest in this interview.

4.3.5 Target Items

A total of 10 idioms (see Table 2) was selected and incorporated into the texts

where they fit the context—5 idioms per text. All target idioms are indicated in the

Collins Cobuild Dictionary of Idioms (2002 edition) as relatively frequent ones,

according to corpus data. They were chosen from a list of 30 idioms (Boers & Yu, in

preparation) that was used to test knowledge of idioms by a previous cohort of

students in the same program in which the current study was conducted. During the

process, I intentionally chose idioms without counterparts in the Chinese repertoire of

standardized figurative expressions. These idioms were shown not to be known yet by

that previous student cohort, and so it was unlikely that they would be known by my

participants. In addition, I asked a colleague, who is an advanced ESL speaker and

has been teaching English in a Canadian public school for two years, whether she was

familiar with any of the 10 idioms. She only knew one of them, and so I was even

more confident that they were unlikely to be known by my participants.

However, as part of the post-test, as an additional check for prior knowledge of

the idioms, the learners were not only asked to recall the meaning of the idioms but

also to state whether they already knew any of them before the experiment. Data

regarding items that were already known were excluded. Even though a pre-test is

arguably a more direct and more reliable means for gauging prior knowledge, the

downside of administering a pre-test is that it will influence learners' subsequent

engagement with the lexical items in the reading text and may make them aware of

the focus of the experiment. Pre-testing knowledge of the meaning of idioms by

means of a free meaning-recall format is also problematic because learners may

proffer a literal interpretation if they understand the words that make up an idiom, but

such a response would not rule out the possibility that the learner also understands the

abstract idiomatic meaning. An alternative might be to use a multiple-choice meaning

recognition test, but this type of test entails a learning opportunity (because the test

taker is exposed on the correct meaning among the test options). For these various

reasons, the option to pretest the target idioms was abandoned in favor of the above

method: (a) selecting idioms from a pre-existing pool, knowledge of which was tested

with a previous student cohort, and (b) asking the participants post-facto if they

already knew the expressions (see Figure 5).

Table 2

Target Idioms and Glosses in the Text

Idiom	Material	Literal-meaning gloss	Figurative-meaning gloss
Throw in the towel	One	in a boxing match, if a coach notices his fighter is losing badly, he may throw a towel into to the ring to signal surrender	give up
Follow suit	One	in a card game, you play a card of the same kind as the previous player	doing the same as what someone else has just done
Wet blanket	One	used to put out a campfire	someone who spoils others' excitement
Take the edge off	One	make a knife or sword less sharp	to reduce the intensity of an unpleasant situation
Stick to your guns	One	on the battlefield, soldiers should stay with their cannons (guns), even if they are under attack	refuse to change your decision
Go against the grain	Two	cutting a piece of wood is harder when you try to do it against its grain (the direction of its fibers)	difficult because it conflicts with established beliefs

Wide of the mark	Two	the mark is the target you aim at in shooting; if you miss badly, your arrow will fall at a wide distance from it	very inaccurate
High and dry	Two	a boat is stuck on land or on a sandbank	in a difficult situation that you cannot do anything about
Put on the back burner	Two	when we cook, we put things to simmer at the back of the stove, so we can use the front burners to attend to other pots or pans	something that does not need immediate attention
Dummy run	Two	in a car crash test, a dummy is used instead of a real person	a test

4.4 Procedures

Most of the participants were recruited from 4 different classes of the MPEd
TESOL Program at the Faculty of Education at Western University. For the sake of
convenience, the main data collection was done at the end of the students' regular on-
site classes of a course, Teaching and Learning Grammar, that was part of the
students' program. It was clarified to the volunteer participants that the study was
unrelated to the course, and the course instructors left the classroom before each data
collection session. I met the participants 3 times or 4 times (4 times in the case of
students who volunteered to be interviewed). In an initial meeting, I introduced the
study and left my email address, so students who wished to participate could email
me. The volunteers were sent the Letter of Information and invited to ask any further
clarification questions. The additional graduate students in Applied Linguistics were
invited to participate via email. Some volunteers said they would not be able to attend
at least one of the sessions in person. For these participants, all data collection was
done in Zoom, but following the same procedures as in the in-person sessions.

The MPEd TESOL cohort consisted of four groups that took their courses in
parallel. The reading activities and post-tests were administered in their existing
groups. Volunteers from one group were assigned to the literal-meaning-only gloss
condition, volunteers from another group were assigned to the abstract-plus-literal-
meaning gloss condition, and volunteers from the remaining two groups were
assigned to the abstract-meaning-only gloss condition. Because the participants

remained in their existing classes, the study must be considered quasi-experimental rather than a pure randomized trial. The additional participants were assigned to treatment conditions depending on when they were available to join the data collection sessions. Owing to scheduling preferences, four MA students and the one PhD student joined the literal-meaning-only group while the other 5 MA participants joined the abstract-plus-literal-meaning group. There were no MA or PhD students in the reading with abstract-only glosses condition as these groups' schedules did not work for them.

 The week after the introductions, the students who volunteered were asked to stay after class for 30-40 minutes to sign the consent form and to do a reading activity with texts with idioms they were unlikely to know. After the reading activity, they were given a post-test to gauge how well they remember the idioms. In the third meeting, one week later, the students were given the delayed post-test (about 20 minutes) again straight after class. After this, volunteers were invited for a follow-up interview (of about 20 minutes) at a time that was convenient for individual volunteers. Once all the scheduled interviews were completed, all the participants received an email with a debriefing form.

 As already mentioned, the three conditions were 1) reading with literal-meaning-only glosses, 2) reading with abstract-meaning-only glosses, and 3) reading with abstract-plus-literal-meaning glosses. The reading task involved two short English articles accompanied by content-related questions. After answering these questions, the participants were tested on their recall of the meaning of the idioms occurring in

the text. The delayed post-test, administered one week later, was intended to gauge long-term retention. As to the follow-up interviews, it was hoped that about 15 students would volunteer (five students from each treatment group). The interview could be considered a stimulated recall interview because the reading texts were used as prompts for the students to talk about how they dealt with the idioms during the reading activity and what helped them to remember them (i.e., to elicit episodic memories of the reading activity and retrieval pathways).

As to the administration of the research instruments (the texts and content-related questions, and the immediate and delayed post-tests), these were printed, and handed out to the participants in the in-person group. For the online sessions, the research instruments were emailed to the participants as PDF files. The participants were not allowed to use resources such as online dictionaries during the reading activities and tests. Figure 5 illustrates the general data collection procedures.

There were no foreseeable ethical issues in this research as students participated on a voluntary basis and all read the same texts. Allocation to a specific reading condition was semi-random, by blind assignment of existing student groups to a condition. At the end of the study, all the participants got a handout with the target idioms and their meanings, and they were briefed about the purpose of the study. A summary of the findings was shared with the students who expressed interest in this. Since the study involves human participation, approval was of course sought from the University's Human Ethics Board (see Appendix P). Only data from students who gave informed written consent were used.

Figure 5

General Data Collection Procedures

Session 1	Phase-1	Introduction of the study and handing out the Letter of Information and Consent Form (LOI/C)		
Session 2	Phase-2	Literal-gloss condition	Abstract-gloss condition	Combination condition
		⬇	⬇	⬇
	Phase-3	Reading Passages + Content-related Questions		
			⬇	
	Phase-4	Immediate Post-test (Idiom Recall Tests)		

⬇

Session 3	Phase-5	Delayed Post-test (With addition of Multiple-choice Test for Idioms)

⬇

Session 4	Phase-6	Interview
		⬇
	Phase-7	Debriefing Form

4.5 Data processing

Participants' responses to the gap-fill test and the meaning recall questions in the immediate post-test and meaning-recall followed by meaning-recognition tests in the delayed post-test were scored dichotomously, 1 point for correct and 0 points for incorrect or missing responses. Scoring was strict. For example, wrongly spelled words in the gap-fill test received no points. It would have been difficult to decide where to draw the line between minor and major spelling mistakes, and in some cases, it would have been difficult to distinguish between spelling mistakes and different words (e.g., *basket* instead of *blanket*). In the meaning-recall tests, no points were for answers that differed significantly from the actual meaning of the idiom. However, some explanations were hard to define as right or wrong because they suggested partial understanding of the target idioms or were somewhat ambiguous. For example, someone's explanation for "*high and dry*" was "people cannot get rid of a situation" but she did not explain what kind of situation. Another participant's answer for "*go against the grain*" was "you are doing something abnormal"; however, "abnormal" is less specific than "conflicting with established beliefs". For the sake of reliability, all the test responses were assessed independently by three raters. Because some participants used Chinese to answer questions, three raters were chosen (two MA students and one PhD student in Applied Linguistics) whose L1 is Chinese. In case of any disagreement among the three raters, the final score was decided by a majority vote. Recall that if participants stated that they already knew a target item before the reading activity, the score on this item was be excluded.

4.6 Data Analysis

All test data were processed by means of Jamovi (Version 2.2, 2021). A generalized mixed model (Gallucci, 2019) was applied to analyze the results of each post-test and to examine whether there was a difference in idiomatic expression acquisition and retention among the three groups who had read texts with literal-meaning glosses, figurative-meaning glosses, or the combination of figurative-plus-literal meaning glosses. We chose the generalized mixed method because, in the current experiment, some data were "missing" (e.g., when a student already knew a certain idiom, and so it had to be excluded for that individual participant). The mixed model assumes that the data are missing at random and therefore does not require complex imputation techniques to replace missing units (Qeunce & van der Bergh, 2004). Compared to the repeated measures ANOVA, the mixed model is a more robust statistical analysis. In case the model indicates a significant difference among the treatment groups, pair-wise comparisons will be applied to examine which reading conditions led to significantly better test scores than others.

The recorded interviews were transcribed and analyzed for the sake of triangulation with the test results. For example, if participants in the literal-meaning gloss condition report their curiosity was piqued about how the gloss information relates to the contextual use of the expressions, then this can help to interpret their test performance, should it point to an advantage for this gloss condition. Moreover, the participants were queried as to whether they expected a test about the glossed items, even though such a test was not announced.

Chapter 5 Results

5.1 Quantitative Data

The descriptive statistics of the test results (gap-fill, meaning recall, and meaning recognition) in the immediate and the delayed post-tests are shown in Table 3. The numbers represent learning gains. If students indicated in the self-evaluation that they had prior knowledge of an idiom, and provided a correct response in the post-tests, then this was not counted as learning gain. Six out of the 37 students (two from each group) reported they already knew one of the 10 idioms before the reading activity. Another 6 students indicated prior knowledge of two target idioms (two from abstract-meaning-only group, one from the literal-meaning-only group and the other three participants were from the combination group). Among the ten idioms, *follow suit* was already known by seven participants before the reading activity; *throw in the towel* by four students; *stick to your guns* by three; *take the edge off* by two, and *wide of mark* and *put on the back burner* were each indicated as already known by only one student.

When we look at the mean score for each idiom among all participants, "*stick to your guns*" got the highest score, with 83.8% participants correctly recalling its meaning it in the immediate post-test and recognizing its meaning in the delayed post-test. The score dropped slightly in the delayed meaning-recall test of the delayed post-test, where 67.6% participants wrote down the correct answer, but it was still the highest among the 10 idioms. In contrast, the scores for "*a wet blanket*", *"take the edge off"* and "*put on the back burner*" were much lower (27%, 29.7% and 29.7%

respectively in the immediate meaning-recall test). In the delayed post-test, only

18.9%, 16.2% and 13.5% of participants wrote down the accurate explanations and

64.9%, 37.8% and 48.6% chose the correct answer in the meaning recognition test.

Table 3 shows large standard deviations for the gap-fill and the two meaning-

recall tests—larger than the average test scores. This indicates that some participants

remembered many of the idioms while others remembered none. The mean scores are

higher for the meaning-recognition test, undoubtedly owing to its multiple-choice

format (allowing for 25% lucky guessing) and the fact that this is a less demanding

test than a recall test. In each section of the post-tests, the learning gains appeared the

best in the reading with abstract-plus-literal glosses group, except in the delayed

meaning-recall test, in which reading with literal-only glosses brought about the better

performance. Meanwhile, reading with abstract-only glosses was associated with the

lowest scores throughout the test sections. These are merely impressions from the

descriptive statistics, however. Inferential statistics are necessary to estimate if any

observed trends are significant.

Four mixed-effects regression models (one per test) were used in the inferential

statistics. Treatment (i.e., gloss condition) was treated as a fixed factor while

participants were considered as a random factor. There was a total of 37 participants,

assigned to one of the three types of treatments: abstract-only glosses ($n= 13$), literal-

only glosses ($n= 12$), and abstract plus literal glosses ($n= 12$).

Table 3

Means and Standard Deviations for Learning Gains of the Target Idioms

Treatment	Gap-fills (IMP)		Meaning Recall (IMP)		Meaning Recall (DLP)		Meaning Recognition (DLP)	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Abstract-only	2.77	4.49	4.54	5.00	2.62	4.41	5.54	4.99
Literal-only	3.50	4.79	4.58	5.00	3.33	4.73	5.75	4.96
Combination	4.00	4.92	4.92	5.02	3.08	4.64	6.00	4.92

Note. IMP= Immediate Post-test. DLP= Delayed Post-test. Each test item is worth 1 point, and the maximum score is 10.

Table 4

The Fixed Effect Omnibus Test of Treatment on the Post-tests

	X^2	*df*	*p*
Gap-fills (IMP)	1.42	2.00	0.491
Meaning Recall (IMP)	0.231	2.00	0.891
Meaning Recall (DLP)	0.768	2.00	0.681
Meaning Recognition (DLP)	0.287	2.00	0.866

The regression models show a substantial role of the random effect, that is, individual differences among the participants. In the gap-fill test, 30% of the variance was associated with this. For the meaning recall in the delayed post-test, 15% of the variance was due to individual differences. For the other two tests (meaning recall in the immediate post-test and the multiple-choice test), the random effect explained 10% and 11% of the variance, respectively.

As for the research question exploring whether there was a difference in the effectiveness among the three reading conditions, the results of the mixed model regression analyses in Table 4 showed that treatment was not a statistically significant factor in any of the four tests *(p= .491, p= .891, p= .681* and *p= .866, respectively).*

Table 5

Comparing Performance on the Post-tests: Estimates of the Fixed Effects on Learning Gains

	Parameter	Effect	Estimate	SE	exp(B)	95% CI for exp(B)		z	p
						Lower	Upper		
Gap-fills	Abstract	2-1	0.411	0.586	1.508	0.478	4.757	0.701	0.483
(IMP)	Literal	3-1	0.688	0.581	1.991	0.637	6.217	1.185	0.236
Meaning	Abstract	2-1	0.0221	0.361	1.022	0.504	2.07	0.0611	0.951
Recall	Literal	3-1	0.1611	0.361	1.175	0.579	2.38	0.4461	0.656
(IMP)									
Meaning	Abstract	2-1	0.363	0.426	1.438	0.624	3.315	0.852	0.394
Recall	Literal	3-1	0.255	0.426	1.290	0.560	2.974	0.598	0.550
(DLP)									

Table 5, continued

	Parameter	Effect	Estimate	*SE*	*exp(B)*	95% CI for *exp(B)*		*z*	*p*
						Lower	Upper		
Meaning	Abstract	2-1	0.0978	0.373	1.10	0.531	2.29	0.262	0.793
Recognition	Literal	3-1	0.1997	0.373	1.22	0.588	2.54	0.536	0.592
(DLP)									

Note. 1= Abstract-only glosses group. 2= Literal-only glosses group. 3= Abstract plus literal glosses group.

5.2 Interview Data

To explore participants' attitudes towards reading with the glosses, the responses in the interview were analysed. A total of 15 participants were interviewed, five from each treatment group. For the question "Did you read the glosses?", 11 students reported looking at all the glosses while reading the texts, but the other four students said they had not paid much attention to the glosses, for different reasons. Two of them (one from abstract-only glosses group and one from the combined-glosses group) thought reading the glosses was too time-consuming as they wanted to finish the texts fast, while the others believed they did not need the glosses to answer the questions about text content. This is reminiscent of previous studies about glossing which found that readers' inclination to inspect glosses depends very much on the task requirements (e.g., Peters et al., 2009).

Rather surprisingly, two students from the reading with abstract-only glosses group said there were expressions they still did not understand (e.g., *throw in the towel*; *wet blanket*) despite having read the texts. In the literal-only glosses group, four students said they still did not get the meaning of some idioms, such as *a wet blanket*, *a dummy run* and *high and dry*. No one from the reading with abstract plus literal meaning glosses group had questions about the idioms' meaning.

Fourteen out of 15 students said the availability of glosses helped their text comprehension. Two from the abstract glosses group thought the glosses helped them understand the text better while all the students from the literal glosses group

mentioned this point in the interview. All interviewees from the combination group

reported that they could visualize the scene evoked by the literal meaning given in the

gloss.

To understand participants' attitudes towards reading with different kinds of

glosses, their responses to an open-ended question in the interview were analysed.

Due to different reading conditions, different questions were asked (see Table 6). The

table also reports students' main reactions and how many times they were mentioned.

Then, participants were invited to rank the 3 text versions from most effective to the

least effective for remembering the expressions. Twelve out of the 15 participants

thought reading with both literal and abstract glosses was the most effective way for

learning idiomatic expressions (Combination glosses> Abstract-only glosses> Literal-

only glosses), reflecting a commonsensical belief that 'more is better'. Only three

participants believed that learners would benefit more from reading with literal-only

glosses (Literal-only glosses> Combination glosses> Abstract-only glosses) because

they thought once learners have made the connection between the abstract and literal

meanings through their own efforts, the memory will last longer. In addition to this,

some participants suggested that different gloss types would be suitable for different

purposes. Participants A6 from abstract-only group and B6 from literal-only group

suggested that reading with abstract meaning glosses (Abstract-only glosses>

Combination glosses> Literal-only glosses) would be the most effective way during a

reading comprehension test. Participant B5 from literal-only group opined that

reading with which kind of glosses should depend on learners' language proficiency: reading with both literal and abstract meaning glosses (Combination glosses > Abstract-only glosses> Literal-only glosses) is suitable for lower language proficiency students as they do not have much knowledge of the cultural background of the target language, and so reading with both abstract and literal meaning glosses may help them to acquire that knowledge. However, those who are at a higher proficiency level should receive abstract-only glosses (Abstract-only glosses> Literal-only glosses> Combination glosses) already have good knowledge about the L2 culture, and so reading with abstract-only glosses directly will suffice.

This experiment was intended to be about incidental vocabulary learning, where incidental learning is operationalized by not announcing the vocabulary tests. Still, students may expect such tests even if they are not forewarned about them and consequently make a deliberate effort to remember the lexical items they meet in a text. I therefore included interview questions to ascertain whether the attested learning gains could really be attributed to incidental learning as a side benefit of a content-focused activity (here, reading for comprehension and evaluation of text content). One question was: "I did not tell you in advance that a test about the idiomatic expressions would follow the reading activities. Did you nonetheless perhaps expect such a test?" Ten out of the 15 students who were interviewed responded "No". Their average scores in all four tests ($M_{IMP\text{-}Gap\text{-}fills}= 0.33$, $M_{IMP\text{-}Meaning\ Recall}= 0.45$, $M_{DLP\text{-}Meaning\ Recall}= 0.29$, $M_{DLP\text{-}Meaning\ recognition}= 0.57$) were lower than those who did expect a follow-up

test ($M_{IMP\text{-}Gap\text{-}fills}$= 0.40, $M_{IMP\text{- }Meaning Recall}$= 0.58, $M_{DLP\text{- }Meaning Recall}$= 0.38, $M_{DLP\text{- }Meaning}$

$_{recognition}$= 0.64). A second question was: "Did you make an effort to commit the

expressions to memory, just in case?". Four students replied "Yes", and three of them

were the ones who also said "Yes" to the previous question. The participants who

gave positive answers tended to perform better in the tests ($M_{IMP\text{-}Gap\text{-}fills}$= 0.33, $M_{IMP\text{-}}$

$_{Meaning Recall}$= 0.57, $M_{DLP\text{- }Meaning Recall}$= 0.33, $M_{DLP\text{- }Meaning recognition}$= 0.67) than those who

said "No" ($M_{IMP\text{-}Gap\text{-}fills}$= 0.33, $M_{IMP\text{- }Meaning Recall}$= 0.44, $M_{DLP\text{- }Meaning Recall}$= 0.29, $M_{DLP\text{-}}$

$_{Meaning recognition}$= 0.56). When they were asked to give reasons, all four students

mentioned their reading habit: they habitually try to hold unfamiliar expressions in

their memory during a reading activity and will intentionally look up the meaning

afterwards. The last question in this part was: "After being tested straight after the

reading activities, did you suspect a delayed test would follow?" Three participants

gave an affirmative response to this question, and explained they expected a delayed

test because of similar research designs they had learned about in their course

readings. They are the ones who also answered "Yes" to the first question. The test

results showed their scores ($M_{DLP\text{- }Meaning Recall}$= 0.43, $M_{DLP\text{- }Meaning recognition}$= 0.70) in the

delayed post-test exceeded those of participants who did not suspect a follow-up

delayed post-test ($M_{DLP\text{- }Meaning Recall}$= 0.29, $M_{DLP\text{- }Meaning recognition}$= 0.57).

Table 6

Participants' Responses to Interview Questions Related to Reading with Different Kinds of Glosses

Treatment Group	Interview Questions	Responses	Number of comments
Abstract-only	The expressions are idioms, that is, figurative phrases that were once used literally. Can you guess their original, literal use?	Yes. I can guess almost all the original meanings.	1
		Not really. I did try during the reading but could not think of much.	4
Literal-only	Did you find it is hard to make connection?	If the picture is easy to imagine it will be easier.	2
		I did not do that because that would take me too much time.	1
		I thought the glosses here are the final meaning, so I did not give it a second thought.	2
	Did it make you feel curious about the current meaning of the expressions?	I spent some time on thinking about the current meaning.	2
		I was in a hurry to finish the content-related questions, so I didn't think carefully.	3
	Do you find information about the origin of the expressions useful?	Literal meaning provided me the bridge to connect the idiom with the context and helped me remember it.	2
		All the idioms are in the key parts of the passage, so the glosses helped me a lot with comprehension.	2
		It didn't help with answering content-related questions, so I didn't pay attention to it.	1

Table 6, continued

Combination	Returning to the information in the glosses about the origin of the expressions, do you find this useful information?	The current meanings had already given me the answer, so I didn't pay attention to the original meaning.	2
		It is good for vocabulary learning but there is too much information.	2
		It is helpful for reading comprehension and long-term memory.	1

In this study, we have found that individual differences account for a

considerable amount of the variance in the post-test performance. The following

excerpts from the interview illustrate some of the differences among individual

participants' experience of the reading activity and the glosses. A stands for Abstract-

only group, B for literal-only group and C for abstract-plus-literal group.

Participant A5 and B5 both misinterpreted *a wet blanket*, but the causes of

confusion were different. Participant A5 made the following statement in the

interview:

> "我看见了注释里面的'spoil'，这是个有很多意思的词嘛，我就把它
>
> 理解成最常见的那一方面了，就是宠坏。放在这个注释里，就是满足某
>
> 人的 excitement 这样的意思了。"["I saw 'spoil' included in the gloss,
>
> which is a word with many meanings, so I interpreted it as the most common
>
> aspect, which is spoiling. Put in this gloss, is to meet someone's
>
> excitement."]

This example illustrates that even glosses intended to give a direct explanation of

the idiomatic meaning can be ambiguous owing to the use of polysemous words in the

definition. That the abstract-only glosses were not crystal clear to all is also illustrated

by participant A12, who felt confused about the meaning of *throw in the towel*:

> "这个意思就是扔毛巾嘛，就有可能是在做家务，之后由于家务太多还
>
> 是什么别的原因，就不想干了，那当然就是把毛巾一扔不继续干了，所
>
> 以就是厌倦了某事，或者因为某件事太麻烦，总是重复性的，就不干

了。""[This means that throwing the towel, it is possible that a person is doing housework, after too much housework or what other reasons, he does not want to continue, then he throws the towel and quits, so it means someone tired of doing something, or because something is too much trouble, so that person is not willing to do it anymore.]

Participant B5 interpreted *a wet blanket* in the following way, based on the literal-only gloss:

"我看到'wet blanket',意思就是湿的毯子嘛,之后用它去扑灭火。意思就是在危急情况的时候,挽救场面,解决问题的那种人。"["I saw 'wet blanket', meaning a wet blanket, after which it was used to put out the fire. It means the kind of person who saves the scene and solves the problem when the situation is critical."].

Another example of this is *dummy run*, which was interpreted differently by participants B5 and B19. Participant B5 reported in the interview,

"我看到'dummy'之后后面提到是用它来取代'真人',在撞车测试嘛。我就觉得这就是电视中总播放的那种,撞车实验。之后用机器去测试数据,看损坏程度,而不是真人靠肉眼去比对。所以我就觉得这个'dummy run'是精准测试的意思了。"["I saw 'dummy', and then the gloss mentioned that it was used to replace 'real people', in the crash test. I think this is the kind of test that is often broadcast on television. After the test, they use a machine to test the data about the degree of damage, rather

than the real person by the naked eye to compare. So, I think this 'dummy

run' is the meaning of precision testing."]

Participant B19, however, reasoned as follows:

"我看到'dummy'想到的就是笨蛋的意思，那么让一个笨蛋去做事

情，肯定就是不靠谱的，结果可能是错的。 而且解释里面也提到了

'test'，所以我就觉得这个习语就是让笨蛋去做测试，这样出来的结

果就是错误的。 " ["When I saw 'dummy' I thought of the meaning of

stupid person, so having a dummy do something must be unreliable and the

result could be wrong. And the explanation also mentioned 'test', so I think

the idiom is to let a dummy do the test, so the result will be wrong."]

Some students from the literal-only glosses group felt confused about some

idioms' exact meanings. Participant B12 was puzzled about *follow suit* while B6 did

not get the meaning of *high and dry*. Participant B12's understanding of *follow suit*

was as follows:

"我觉得'follow suit'，就是'follow'一些事情吧，或者规则和条

款。之后又看到'suit'，就想到是套装。很像一些晚宴会有'dress

code'一样，别人穿什么，你也要穿什么。比如，晚宴的主题是红色，

那你就要穿红色的衣服，需要遵守人家的要求。但是，给出来的原始意

义，有提到说是 card game，我就不是很能把'suit'和'card game'联

系起来。他们之间是有什么意思。但是我能大概猜到这个 idiom 是要说

学别人怎么做你就怎么做的意思。 " ["I think 'follow suit' is 'follow' some

things, or rules and terms. Then I saw 'suit' and thought it would refer to

clothes. It's very much like some dinner parties where there is a 'dressing

code' and you have to wear whatever others are wearing. For example, if the

theme of the dinner is red, then you have to comply with people's

requirements and wear red. However, the original meaning given here, has

mentioned that it is in a card game. So, I got confused, I cannot connect 'suit'

with 'card game'. What is the meaning between them? But I can probably

guess that this idiom is to say doing the same as what someone else has

done.”]

Even though participant B12 figured out the figurative meaning of *follow suit* in her

own way and recalled it in the meaning recall of the immediate post-test, she failed to

produce the correct answer in the delayed post-test.

In a similar vein, participant B6 felt confused about the literal underpinning

proposed for the expression *high and dry*:

“我不太懂‘high and dry’的意思，字面的意思就是又干又高。之后就

是看到给出的解释，说一条船困在了沙滩上。但是海滩并不是很高的地

方啊，所以我就很不能理解，这个 idiom 到底是什么意思。” [“I don't

quite understand the meaning of 'high and dry'. I know it literally means a

place which is dry and high. Then I saw the explanation given is about a boat

was stuck on the beach. But the beach is not a very high place, so I cannot

get the meaning of this idiom.”]

B6 got one point in the meaning recall of immediate post-test, but she did neither recall nor recognize the meaning in the delayed post-test.

As one of the examples above already illustrated, it is worth emphasizing that the gloss type did not determine the way participants engaged with the idiomatic meanings. For example, some students in the abstract-only gloss condition spontaneously referred to imagery. Participant A5, who got the highest scores on the four tests in the group, mentioned that it is easy to picture a scene for *throw in the towel*, *go against the grain*, *wide of mark*, *high and dry* and *dummy run*.

"这些 idiom 还是很好想象他们的意思的就比如，'throw in the towel'就很明显呀，投降的时候会扔出来白手绢。'go against the grain'就是一些东西它都是有自己的纤维走向的嘛，之后你不顺着它，你偏要逆着他的纹理来，那肯定就要困难一些。还有'wide of mark'就很好理解，mark 就是你的目标，就像是你扔飞镖时候的圆心一样，之后你扔的很偏，就距离圆心很远。这不就是错的离谱，差得远的意思嘛。还有就是'high and dry'就是被困在了一个又高又干的地方，就相当于陷入了很无助的境地，什么也做不了也改变不了的一种地步。最后还有就是'dummy run'，因为 dummy，我之前知道有假人的意思，就相当于是说用一个假人去做一些比较危险的测试，比如车辆出厂之前都是要做那种碰撞试验之类的，就会放一个假人进去，而不是真人。这几个就是让我觉得还是很好理解的，就是看一眼就能出现那个画面，很好联想到意思。"["These idioms are very easy to imagine the meanings of. For

example, 'throw in the towel' is very obvious, it refers to the time of

surrender and people will throw out the white handkerchief. 'Go against the

grain' means something that has its own fiber directions. If you do not follow

it, which means you go against his grain, that thing will become difficult.

And 'wide of mark' is easy to understand, the mark is your target, just like the

center of the circle when you throw a dart. After you are throwing very far

from the center of the circle, you are far from your goals. That's what it

means to be so wrong. 'High and dry' refers to something trapped in a high

and dry place, which is equivalent to being in a very helpless situation,

nothing can be done, and the situation cannot be changed. Finally, there is

'dummy run', because I know the meaning of dummy, it means a fake human.

It could mean that a dummy has been used to do some dangerous tests, such

as vehicles need to do crash tests before leaving the factory and so on. A

dummy will be used instead of a real person. These are the ones that make

me think it is imaginable and easy to get the meaning."]

The test results of this participant show that all the idioms she mentioned here were

correct in the meaning recall and multiple-choice test. However, the mental imagery

of the concrete scenes she associated with the idioms did not help the participant

much in the form-recall test.

Participant A14, who was the second-best learner in the immediate post-test from

the same group, reported that by reading the explanation of *stick to your guns* and *put*

something on the back burner she was able to picture the meaning of these

expressions:

> "'Stick to your guns'就一下子就有那个画面了。'stick'就是粘，沾
>
> 住的意思嘛。那你跟枪粘在一起不就是士兵需要每时每刻人在枪在的意
>
> 思吗。就感觉是坚守住自己的职业底线。放在这个文章里，就是坚守自
>
> 己的立场和底线。之后就是'put something on the back burner'，可能
>
> 是因为我在家经常做饭吧，就特别能感同身受，把什么东西放在一边的
>
> 灶眼儿上，就肯定是需要长时间炖煮的东西，因为最常用的灶眼需要用
>
> 来炒菜之类的。之后再结合一下实际情况，国外的灶眼排列都是前后这
>
> 样的，不像国内是左右排列。所以就很好理解，放在后面的灶眼上，就
>
> 肯定是不着急做好的东西啦。"["For 'stick to your guns', I have that
>
> image at once. 'Stick' is sticky, the meaning of staying together with
>
> something. So, stick to your guns is like a soldier needs to be together with
>
> his gun all the time. It feels like holding on to the bottom line of one's
>
> profession. Put in this context, is to adhere to their own position and the
>
> bottom line. After that is 'put something on the back burner', probably
>
> because I often cook at home, it is particularly empathetic, put something on
>
> the side of the stove, it must be something that needs to be stewed for a long
>
> time, because the most commonly used stove needs to be used for stir-fry and
>
> so on. After that, combined with the actual situation, foreign cookers are
>
> arranged in such a way that the front and back, unlike the way in China

which is left and right. So, it is very understandable, put on the back of the

stove, it is certainly not in a hurry to do things."]

She wrote down the correct answers for both idioms throughout the four tests. For the

other idioms, she was less successful, only getting three more points in the meaning

recall of immediate post-test and no extra points in the delayed post-test.

Mental imagery was expected far more in the gloss conditions that presented the

students with the original, literal use of the expressions. Even so, some participants

were more likely than others to refer to images. Participant B12 who is the best

learner from the literal-only glosses group said she found the glossary of *stick to your*

guns and *high and dry* very helpful. They were easy to interpret and thus it is simple

to imagine:

> "这几个 idiom 是我看完旁边的注释之后，让我恍然大悟的。比如说，
>
> 'stick to your guns'，就是士兵必须要跟他们的 cannons 呆在一起，即使
>
> 是在被袭击的情况下。这不就是坚守住自己的岗位。我看到这个注释，
>
> 就一下子能联系到这个 idiom 是在说什么了。另外一个就是 'high and
>
> dry'，给出的注释里面说是一个小船被卡在了沙子里，那肯定啊。本
>
> 来船是要在水里才能行进的，卡在了沙子里就是根本动不了，什么也做
>
> 不了，哪也去不了。就能感觉到是陷入了很无助的境地。" ["These are
>
> a few idioms that dawned on me after I read the glosses next to them. For
>
> example, 'stick to your guns', is about soldiers who have to stay with their
>
> cannons, even when they are under attack. It's just like holding on to your

post. When I saw this note, I could instantly connect to what this idiom was

talking about. Another one is 'high and dry', the glossary given says that a

small boat is stuck in the sand. Originally a boat should be in the water to

move. However, it is stuck in the sand, so the boat cannot move. I can

understand this means to be caught in a very helpless situation"] This

participant got both items correct in the meaning recall immediate post-test

and multiple-choice test. The imagery did not help the participant in the

form-recall test, however.

  Participant C3, the best learner from reading with abstract-plus-literal meaning

group expressed the following idiomatic expressions evoked images: *a wet blanket*,

go against the grain, *high and dry*, *dummy run*. Full scores were obtained for these

idioms in the four tests. She described the process of learning these idioms as

follows.

> "我有仔细的看旁边所有的注释。因为是分为两部分，我记得一个是他
>
> 现在的含义，一个是他之前的意思，相当于是背景故事那种，我看着还
>
> 是觉得挺有意思的。比如说'wet blanket'，原始注释里说是浇灭篝火
>
> 的毯子，再看现在的意思有扫兴的意思。我就能明白这个篝火是说的快
>
> 乐之火，那作为湿毯子这个人就是会毁坏别人快乐的人，很扫兴。还有
>
> 'go against the grain'，在注释里面我就找到了同义词。现在的意思有
>
> 一个'conflict'，下面的原始意思里面有一个'against'，那他们后面
>
> 跟着个词就是相近意思。Grain 就相当于是 beliefs。一下子就明白了意

思。违背了原始的初衷。下一个是'high and dry'一艘小船被困在了沙子里，之后现在意义说是一种很困难的处境，却什么都做不了。这样就很好理解，船困在沙子里这种无助的处境了。最后一个印象深刻的是'dummy run'，用 dummy 代替真人去进行撞车试验。再放在原文中就是我们要认真对待环境问题。"["I have carefully read all the glosses next to the text. It (each glossary) includes two parts, I remember one is the current meaning, one is the original meaning, equivalent to the kind of backstory, I find it quite interesting to look at. For example, 'wet blanket', the original meaning says it is a blanket to put out the campfire, and then look at the current meaning has the meaning of spoilers. I can understand that this campfire is said to be the fire of happiness, that as a wet blanket this person is the person who will ruin the happiness of others, very spoiled. And 'go against the grain', I found the synonym in the glossary. Grain is equivalent to beliefs. So, it means something goes against the original intention. The next one is 'high and dry', a small boat is trapped in the sand, and the current meaning is in a very difficult situation, but nothing can be done. So, this is about a helpless situation of a boat trapped in the sand. The last one that impressed me is 'dummy run', using dummy instead of real people to carry out crash tests. Then put in the original text means we need to treat environmental problem seriously."]

From her description, we find that providing students with both literal and abstract meaning may help them to find the connection and retain the meanings in long-term memory.

For the last interview question, "Language courses sometimes include texts accompanied by glosses to clarify vocabulary. Is this something you remember from your own language learning or teaching experience?" Only four participants reported that they had read glossed texts before. However, all those texts were in the after-class reading books published by English speaking countries.

Chapter 6 Discussion

In this chapter, several findings will be discussed by linking the results of the statistical analysis with the responses collected in the participants' interviews.

Referring to the research question of this study, that is, whether there is a difference in the effectiveness of the three types of glosses, I hypothesized that literal-only glosses would lead to better post-test performance than abstract-only glosses. This hypothesis was based on Boers (2000). I refrained from formulating a hypothesis regarding the abstract-plus-literal glosses, because this treatment condition was missing in the precursor research. No significant differences among the three groups were found in either the immediate post-tests or the delayed post-tests. Going by the descriptive statistics, reading with abstract-plus-literal meaning glosses was the most beneficial except in the meaning recall of the delayed post-test. The literal-only glosses appeared to be second most effective, overall. This may be because students in this condition needed to invest some cognitive effort into making a connection between the literal meaning and the contextualized current meaning of the idioms (although the interview data reveal that not all the students engaged in this elaborative processing). The investment of cognitive effort typically benefits long-term retention (*cf.* the *Desirable Difficulties* framework, e.g., Bjork, 1994) and so this may explain why the literal-only glosses led to slightly better performance in the delayed meaning recall test compared to the combined glosses, which provided all the information. However, in the meaning recognition test (i.e., the multiple-choice test), which is a

less demanding test, the literal-only group did not outperform the abstract-plus-literal

glosses group. It is worth mentioning that the meaning recognition test probably

advantaged the students who had seen the definitions of the idiomatic meanings in

their glosses, since these were the definitions presented in the multiple-choice items

and so it was a matter of recognizing them. However, the students who had seen only

the literal-meaning glosses may have found it harder to match their own mental

representation of the idiomatic meanings to the definitions used in the multiple-choice

test.

It is worth reiterating that individual differences accounted for a considerable

amount of the variance in post-test performance. The examples of participants A5 and

B5's misunderstanding to *a wet blanket* as well as participants B5 and B19's

interpretation to *dummy run* illustrate that there were differences regarding which of

the idioms individual participants found difficult to understand, and this was not

always associated with their gloss condition. Participant B5's answer to *wet blanket*

illustrates that using hints about the literal underpinning of an idiom does not always

lead to an appropriate interpretation of the idiomatic meaning, even if this information

is offered in combination with discourse context that is expected to guide the

interpretation.

Participants from all three groups mentioned the benefits of imagery for learning

idioms. The example of Participant A12 explaining *throw in the towel* illustrates that

this participant from the abstract-only-glosses group spontaneously tried to connect

the figurative meaning of the idiom to a concrete scene in which the expression is

used in a literal sense. There is indeed evidence from psycholinguistics experiments

that L2 learners often activate both literal and figurative readings of idioms, even

when they are not prompted to do so (e.g., Cieślicka, 2006).

The example of Participant B12's explanation of *follow suit* illustrates that

learners may find the given hint about the literal underpinning too obscure and then

choose to connect the figurative meaning of an idiom to a literal origin they find more

plausible or easier to understand, even if the latter is not "etymologically" accurate.

We found from relating students' interview responses to their test responses that

the imageability of idioms may influence meaning recall but it does not seem to

influence form recall as much. This illustrates that "semantic elaboration" (in this case

connecting figurative meanings to literal underpinnings) is beneficial for meaning

retention but cannot always be expected to be equally beneficial for form retention

(Barcroft, 2015). This is understandable, because the form-recall (gap-fill) test

required recall of specific words, such as *towel*, *grain*, and *mark*, while a mental

picture of a scene does not necessarily foster memories of these precise words.

Instead, target words were occasionally substituted by synonyms, such as *target*

instead of *mark* for the idiom *wide of the mark*. There was also some evidence of

cross-item interference, for example when instead of *guns* a participant wrote *line* to

complete *stick to your _____*, likely owing to interference from *toe the line*.

Participants would also write down words that occurred in the proximity of the idiom

in the text or words from the definition provided in the gloss, such as *rules* to complete *stick to your _____*. This is reminiscent of earlier research which suggested that elucidations of meaning are beneficial for meaning retention, but not necessarily for learners' retention of the precise lexical makeup of idioms (e.g., Boers et al., 2009). In short, the acquisition and retention of the idioms' form (i.e., their precise lexical composition) does not seem to have a strong relation with whether the meaning of the expression is easy to "picture".

The interviews also revealed that some participants tackled the reading activity as they would a time-pressured reading comprehension test, even though they had not been instructed to do so. The amount of time that learners think they need to invest to match a reading purpose will also influence what use they will make of glosses and what kind of glosses they find suitable. If students are under no time pressure, then they may find glosses which provide information beyond the contextual meaning of words or phrases suitable. In those circumstances, L2 instructors can then also encourage their students to read with literal-only glosses and try to infer the figurative meanings. If students tackle a reading activity under time pressure, however, then they will find concise glosses that give direct access to the contextual meanings more appropriate. That some participants felt they needed to get the reading task done as fast as possible was yet another variable that I had not anticipated, and which emerged thanks to adopting a mixed methods approach.

Chapter 7 Conclusions and Implications

This research project aimed to examine the effect of different support methods (literal and/or figurative glosses) on high intermediate to advanced ESL learners' vocabulary acquisition from reading. There is very little research on glossing that focuses on MWIs, not to mention idioms. One early study which included idioms among the target items was Boers (2000), in which the effectiveness of literal-meaning glosses and abstract-meaning glosses was compared, but this study had shortcomings. For one, it did not explore the effect of glosses that present both abstract plus literal meanings. Being exclusively a pen-and-paper experiment, it did not explore how the learners engaged with the glosses either. Another limitation was that the test was a contextualized form-recall test, without ascertaining that the learners really understood the target expressions. Thus, I decided to conduct a mixed methods study, comprising a classroom experiment and interviews. In this research, both quantitative and qualitative data have been collected and analysed. Three reading conditions have been employed: reading with literal-only glosses, reading with abstract-only glosses, and reading with both abstract plus literal glosses. Originally, I had planned to include a reading condition without glosses, to obtain "baseline" data of how much learning happens without such support and to see if glossing truly helps. Owing to challenges with recruiting sufficient participants, I decided to let go of this no-gloss condition. There already is ample research indicating that reading with glosses is beneficial for vocabulary learning relative to reading without glosses

(Boers, 2022; Yanagisawa et al., 2020, for reviews), and so I decided to give priority to comparing different glosses rather than establishing once more that glossing per se is helpful.

Even though we cannot tell if students will benefit more from one than another type of glosses since there were no statistically significant differences detected among the three reading condition groups, the findings from the descriptive data and the interviews suggest reading with both literal and abstract glosses may be the most effective way to retain the meaning of idioms in memory, at least in the short term. In addition, the acquisition of an idiom's meaning is related to whether the idiom is easy to imagine and whether the explanation is transparent for students.

The interview responses suggest that learners find glossed reading materials useful. The reasons why participants find it helpful varied, however. Some like glosses primarily because they make the reading experience more fluent by eliminating the need to look items up in a dictionary (46.7%), while others see them first and foremost as a good opportunity to learn new words through reading (53.3%).

This research illustrates the usefulness of mixed methods research, where qualitative data shed light on quantitative data. Based on what we found in the interview, it is naïve for researchers to assume learners will process the information or go about the given tasks according to the researchers' expectations. Even though a learning condition is intended to gauge incidental learning, students may spontaneously switch to intentional learning. Although certain information (e.g.,

glosses directly explaining idiomatic meaning) is meant to be totally transparent, this is not always the case from the student's perspective. Even though one treatment is meant to stimulate mental operations that are different from another treatment (e.g., visualization of idiomatic meanings), what mental operations students really perform in their learning conditions cannot be controlled (for example, even in the abstract-only glosses condition, some learners spontaneously evoked images of literal uses of the expressions).

The statistics revealed a substantial role for individual differences, but without the interviews, I would not have been able to evaluate the nature of those differences in the present study. One explanation for some participants' comparatively high post-test scores was that they tackled the reading activity as a vocabulary-learning task (i.e., deliberate learning) instead of a "mere" text comprehension task (with vocabulary picked up incidentally). These participants were then also likely to expect a vocabulary test. It is possible that the nature of the glosses puts learners into a deliberate learning mode as well. A brief clarification of the meaning of a word or a phrase as it is used in the text is likely to be interpreted as mere support for text comprehension, but an elaborate gloss that includes information which appears not essential for text comprehension may be taken as a signal that the activity is not just a text comprehension task.

The interview data also illustrated the role of learning style or cognitive style differences. Regardless of the gloss condition, some participants were more likely

than others to (try to) relate the contemporary, figurative meanings of the idioms to

literal readings evoking images of concrete scenes. This is reminiscent of studies

which found correlations between L2 learners' retention of deliberately studied

idioms and their habitual use of mental imagery according to a cognitive-style

questionnaire (e.g., Boers et al., 2006). For learners who will spontaneously resort to

mental imagery during learning tasks, hints about the literal underpinnings of idioms

may possibly be redundant in cases where those underpinnings are relatively

straightforward to these learners. On the other hand, learners who are less inclined to

spontaneously conjure up images when learning idioms may benefit more from such

hints and may also be helped by actual pictures or drawings representing the original,

literal use of the expressions (Boers et al., 2009).

The findings of this study have pedagogical value. Reading glossed texts in the

L2 class seems not to be common, according to textbook analyses (Boers, 2022) and

according to statements made by interviewees in this study. And yet, it has been

shown in many studies that glossing facilitates reading comprehension and lexical

acquisition. During the interview, 11 out of 15 students reported that they had never

met glosses in their EFL courses, and the remaining four participants expressed they

had only encountered glossed texts in their after-class reading books. Besides, the

glossing they remembered concerned single words, not idioms.

The prediction based on Boers (2000) was that literal-only glosses would bring

about comparatively good learning, but the findings only partly confirm this. One

possible reason is that it depends on the individual idioms: This gloss type can only "work" if students find the information about the literal underpinning helpful enough to figure out the actual idiomatic meaning. Both my quantitative and qualitative data suggest it is very difficult for a teacher/researcher to rely on intuition in this regard. I was rather confident that the combination of literal underpinnings and the supporting discourse context in which the idioms occurred would help my highly proficient participants to work out the idiomatic meanings, but the interviews revealed this was not always the case. To go beyond teacher intuition, it would be useful to collect data from the student population for whom one is creating L2 reading materials as to whether they are likely to find the given information helpful. This could be done, for example, by compiling a list of relatively frequent (and thus useful) idioms and eliciting student responses for these. Students could be given both the literal and the figurative meanings of idioms and asked to rank the idioms based on how transparent they judge the connection to be (see, Wang et al., 2020 for such a procedure). If the connection is felt to be relatively clear, then the literal underpinning could be used to engage the next student cohorts in an inferencing activity. If the connection is considered opaque, then it is best to directly explain the figurative use, and only then add the literal underpinning if this can make the idiom easier to remember. Adapting gloss type to the nature of the idioms could also be done in the form of action research, where individual teachers improve their materials regularly, based on how effective (or ineffective) they were with a previous student cohort. For example, based

on the test result of the current study, 83.8% of all the participants (100% of the

literal-only gloss group) correctly recalled the meaning of "*stick to your guns*". For

the idiom "*take the edge off*", however, only 29.7% of the students (15.4% of the

literal-only gloss group) wrote down the correct explanation, so a different type of

gloss is needed for this idiom in which the figurative meaning is explicitly clarified.

Therefore, in the future, I may use a literal-only gloss to teach "*stick to your guns*",

since this seems to suffice, and provide students a literal-plus-abstract meaning gloss

for "*take the edge off*" (see Figure 6).

Figure 6

Example of Teaching Different Idioms with Different Kind of Glosses

It is understandable if you give in because no parent wants to be a wet blanket. Still, it is better to remove yourself and your child from the situation that triggered the nagging and go to a neutral spot without such outside interference. This will **take the edge off** the situation, and once the child has calmed down, you should reward him or her with praise for calming down. Explain to him again what your previously established rule concerning nagging and whining is. Now you are reinforcing the appropriate behavior that you want to see happen. **Sticking to your guns** is extremely important.

Take the edge off: (origin) make a knife or sword less sharp; (current) to reduce the intensity of an unpleasant situation.

Stick to your guns: (origin) on the battlefield, soldiers should stay with their cannons (guns), even if they are under attack.

Chapter 8 Limitations and Directions for Future Research

This research has several major limitations. The first limitation that needs to acknowledged is the small sample size, both in terms of participants (N=37) and target idioms (N=10). Even though 1,480 observations have been included in the mixed-effect regression model, this may yet be insufficient to detect statistically significant difference in the effectiveness of reading with different kinds of glosses. We could consider the findings of the current study as exploratory, and future replication studies are highly welcome. For instance, the present study did not include a group who read the articles without any glosses; therefore, it is impossible to measure the effect of the presence of glosses as such on incidental idiom acquisition. We initially designed four treatment groups (reading with literal-only glosses, abstract-only glosses; abstract plus literal glosses and non-glosses). The non-glosses group was planned to be the control group, and it could have helped us to see if students managed to figure out the meaning of the idioms from context, which would then put the benefits of the glosses into perspective. However, we had to exclude the non-glosses group because we had too few participants group to be distributed across four treatment conditions. In addition, as mentioned previously, several studies have already furnished compelling evidence that reading with glosses is more effective than without glosses, and so I felt it justified to let go of the non-gloss condition and to prioritize the comparison of the three different glossed-reading conditions.

Three more limitations concern the research instruments. First, we originally planned to have all tests conducted in-person. Because of the scheduling problem of some participants, we had to change one of the treatment groups into an online format. However, we did not prepare any online testing software, and so we had to send the PDF files to all the online group participants and hope they would follow the instructions not to re-read the articles as they took the post-tests. In future research, all participants should complete the test under the same circumstances, either in-person in a classroom or online with a prepared digital version of the materials.

Second, for reading with both literal and abstract glosses group, we presented the current (abstract) meaning before the original (literal) meaning (see Figure 7). Some students may just have read the current meaning and ignored the origin to save time; therefore, the connection between literal and abstract meaning we expected may not always have been established. We presented the information in this order, simply because this is also what is done in idiom dictionaries which include "etymological" notes. In future replications, it may be worth including a reading condition with glosses that present the literal underpinning before the current, figurative meaning of the expressions. This could make it less likely that the information about the literal origins is overlooked. When it comes to finding out what information learners pay attention to, the present study relied on retrospective interviews. A more "sophisticated" way of assessing learners' allocation of attentional resources during reading would be to use eye-tracking technology (e.g., Warren et al., 2018).

Figure 7

Example of Reading with Combination Glosses

Children use pester power to persuade parents to buy them things that they don't really need. The nag factor not only gets parents to buy things but is also used to manipulate parents to do things they normally would not. A 2002 survey found that most children repeatedly request an item until their parent **throws in the towel.** Some will ask more than 50 times for a particular product.

Throw in the towel: (current) give up; (origin) in a boxing match, if a coach notices his fighter is losing badly, he may throw a towel into to the ring to signal surrender.

Moreover, we only conducted one form-recall test as immediate post-test in the current study. Had we administered another form-recall test in the delayed post-test, it might perhaps have revealed more differences in learning gains.

In the future research, it will be worth exploring the possibility of using other kinds of glosses to learn idioms. For example, using actual pictures of the literal meanings. In a previous study (Boers et al., 2009), researchers applied pictorial glosses in the deliberate teaching of idioms and found this did not benefit learners' retention of form (i.e., the lexical composition) much (possibly because the pictures drew more attention than the words), but it did seem to facilitate meaning-recall. Another possibility is if all the participants share the same L1, it is worth examining whether reading with L1 glosses leads to better retention in memory as the use of L1 glosses makes the information easier to process.

References

Acha, J. (2009). The effectiveness of multimedia programmes in children's

    vocabulary learning. *British Journal of Educational Technology*, *40*, 23–31.

    https://doi.org/10.1111/j.1467-8535.2007.00800.x

Altenberg, B., & Granger, S. (2001). The grammatical and lexical patterning of make

    in native and non-native student writing. *Applied Linguistics*, *22*, 173–195.

    https://doi-org.proxy1.lib.uwo.ca/10.1093/applin/22.2.173

Barcroft, J. (2015). *Lexical input processing and vocabulary learning*. John

    Benjamins.

Barfield, A., & Gyllstad, H. (Eds.) (2009). *Researching collocations in another*

    language: Multiple perspectives. Palgrave Macmillan.

Bensoussan, M., & Laufer, B. (1984). Lexical guessing in context in EFL reading

    comprehension. *Journal of Research in Reading*, *7*(1), 15–32.

    https://doi.org/10.1111/j.1467-9817.1984.tb00252.x

Bestgen, Y. (2017). Beyond single-word measures: L2 writing assessment, lexical

    richness and formulaic competence. *System*, *69*, 65–78.

    https://doi.org/10.1016/j.system.2017.08.004

Bjork, R. A. (1994). Memory and metamemory considerations in the training of

    human beings. In J. Metcalfe, & A. Shimamura (Eds.), *Metacognition:*

    Knowing about knowing (pp. 185–205). Massachusetts Institute of Technology

    Press.

Boers, F. (2000). Enhancing metaphoric awareness in specialised reading. *English for Specific Purposes*, *19*, 137–147.

https://doi.org/10.1016/S0889-4906(98)00017-9

Boers, F. (2020). Factors affecting the learning of multiword items. In S. Webb (Ed.), *The Routledge handbook of vocabulary studies* (pp. 143–157). Routledge.

https://doi.org/10.4324/9780429291586-10

Boers, F. (2022). Glossing and vocabulary learning. *Language Teaching*, *55*(1), 1−23.

https://doi.org/10.1017/S0261444821000252

Boers, F., Eyckmans, J., & Stengers, H. (2006). Motivating multiword units: Rationale, mnemonic benefits, and cognitive style variables. In S. H. Foster-Cohen, M. M. Krajnovic, & J. M. Djigunovic (Eds.), *EUROSLA Yearbook 6* (pp. 169–190). Benjamins.

Boers, F., Eyckmans, J., Kappel, J., Stengers, H., & Demecheleer, M. (2006). Formulaic sequences and perceived oral proficiency: Putting a lexical approach to the test. *Language Teaching Research*, *10*, 245–261.

https://doi.org/10.1191/1362168806lr195oa

Boers, F., & Lindstromberg, S. (2009). *Optimizing a lexical approach to instructed second language acquisition*. Palgrave Macmillan.

Boers, F., & Lindstromberg, S. (2012). Experimental and intervention studies on formulaic sequences in a second language. *Annual Review of Applied Linguistics*, *32*, 83–110.

https://doi.org/10.1017/S0267190512000050

Boers, F., Piquer Píriz, A. M., Stengers, H., & Eyckmans, J. (2009). Does pictorial

elucidation foster recollection of figurative idioms? *Language Teaching

Research*, *13*, 367–388.

Boers, F., Warren, P., He, L., & Deconinck, J. (2017). Does adding pictures to glosses

enhance vocabulary uptake from reading? *System*, *66*, 113–129.

https://doi.org/10.1016/j.system.2017.03.017

Boers, F., & Webb, S. (2015). Gauging the semantic transparency of idioms: Do

natives and learners see eye to eye? In R. Heredia & A. Cieślicka (Eds.),

Bilingual figurative language processing (pp. 368–392). Cambridge

University Press.

Boers, F., & Yu, X. (in preparation). Comparing ways of teaching L2 idioms: The

roles of contexts, literal underpinnings, and successful inferencing.

Bui, T., Boers, F., & Coxhead, A. (2020). Extracting multiword expressions from

texts with the aid of online resources. *International Journal of Applied

Linguistics*, *171*, 221–252.

https://doi.org/10.1075/itl.18033.bui

Cheng, Y-H., & Good, R. L. (2009). L1 glosses: Effects on EFL learners' reading

comprehension and vocabulary retention. *Reading in a Foreign Language*, *21*,

119–142.

Cieślicka, A. (2006). Literal salience in online processing of idiomatic expressions by

second language learners. *Second Language Research*, *22*(2), 115–144.

https://doi.org/10.1191/0267658306sr263oa

Cieślicka, A. B. (2012). Do nonnative language speakers chew the fat and spill the

     beans with different brain hemispheres? Investigating idiom decomposability

     with the divided visual field paradigm. *Journal of Psycholinguistic*

     Research, *42*, 475–503.

     https://doi.org/10.1007/s10936-012-9232-4

Cobb, T. (2018). From corpus to CALL: The use of technology in teaching and

     learning formulaic language. In A. Siyanova-Chanturia & A. Pellicer-Sánchez,

     A. (Eds.), *Understanding formulaic language: A second language acquisition*

     perspective (pp. 192–210). Routledge.

Conklin, K., & Schmitt, N. (2012). The processing of formulaic language. *Annual*

     Review of Applied Linguistics, *32*, 45–61.

     https://doi.org/10.1017/S0267190512000074

Cooper, T. C. (1999). Processing of Idioms by L2 Learners of English. *TESOL*

     Quarterly, *33*(2), 233–262. https://doi.org/10.2307/3587719

Crossley, S.A., Salisbury, T., & McNamara, D.S. (2015). Assessing lexical

     proficiency using analytic ratings: A case for collocation accuracy. *Applied*

     Linguistics, *36*, 570–590. https://doi.org/10.1093/applin/amt056

Davis, J. N. (1989). Facilitating effects of marginal glosses on foreign language

     reading. *The Modern Language Journal*, *73*, 41–48.

     https://doi.org/10.1111/j.1540-4781.1989.tb05308.x

Eyckmans, J., Boers, F., & Stengers, H. (2007). Identifying chunks: Who can see the

     wood for the trees? *Language Forum*, *33*, 85–100.

Foster, P. (2001). Rules and routines: A consideration of their role in the task-based

    language production of native and non-native speakers. In M. Bygate, P.

    Skehan & M. Swain (Eds.), *Researching pedagogic tasks: Second language*

    learning, teaching, and testing (pp. 75–93). Longman.

Gallucci, M. (2019). *GAMLj: General analyses for linear models.* [jamovi module].

    Retrieved from https://gamlj.github.io/.

Godfroid, A., Boers, F., & Housen, A. (2013). An eye for words: Gauging the role of

    attention in L2 vocabulary acquisition by means of eye-tracking. *Studies in*

    Second Language Acquisition, *35*, 483–517.

    http://doi-org.proxy1.lib.uwo.ca/10.1017/S0272263113000119

Gettys, S., Imhof, L. A., & Kautz, J. O. (2001). Computer-assisted reading: The effect

    of glossing format on comprehension and vocabulary retention. *Foreign*

    Language Annals, *34*, 91–99.

    https://doi.org/10.1111/j.1944-9720.2001.tb02815.x

Granger, S. (1998). Prefabricated patterns in advanced EFL writing: Collocations and

    formulae. In A.P. Cowie (Ed.), *Phraseology: Theory, analysis, and*

    applications (pp. 145–160). Oxford University Press.

Hoang, H., & Boers, F. (2018). Gauging the association of EFL learners' writing

    proficiency and their use of metaphorical language. *System*, *74*, 1–8.

    https://doi.org/10.1016/j.system.2018.02.004

Holley, F. M., & King, J. K. (1971). Vocabulary glosses in foreign language reading

    materials. *Language Learning*, *21*, 213–219.

https://doi.org/10.1111/j.1467-1770.1971.tb00060.x

Hulstijn, J. H. (1992). Retention of inferred and given word meanings: Experiments in

incidental vocabulary learning. In P. J. L. Arnaud & H. Bejoint (Eds.),

Vocabulary and applied linguistics (pp. 113–125). Palgrave Macmillan.

Jacobs, G. M. (1994). What lurks in the margin: Use of vocabulary glosses as a

strategy in second language reading. *Issues in Applied Linguistics*, *5*, 115–137.

Jacobs, G. M., Dufon, P., & Hong, F. C. (1994). L1 and L2 vocabulary glosses in L2

reading passages: Their effectiveness for increasing comprehension and

vocabulary knowledge. *Journal of Research in Reading*, *17*, 19–28.

https://doi.org/10.1111/j.1467-9817.1994.tb00049.x

Ko, M. H. (2005). Glosses, comprehension, and strategy use. *Reading in a Foreign

Language*, *17*, 125–143.

Kost, C. R., Foss, P., & Lenzini, J. J. (1999). Textual and pictorial glosses:

Effectiveness on incidental vocabulary growth when reading in a foreign

language. *Foreign Language Annals*, *32*, 89–97.

https://doi.org/10.1111/j.1944-9720.1999.tb02378.x

Krashen, S. (1989). We acquire vocabulary and spelling by reading: Additional

evidence for the Input Hypothesis. *The Modern Language Journal*, *73*, 440–

464. https://doi.org/10.1111/j.1540-4781.1989.tb05325.x

Laufer, B. (1997). The lexical plight in second language reading: Words you don't

know, words you think you know, and words you can't guess. In J. Coady &

T. Huckin (Eds.), *Second language vocabulary acquisition: A rationale for pedagogy* (pp. 20–34). Cambridge University Press.

Laufer, B., & Waldman, T. (2011). Verb-noun collocations in second language writing: A corpus analysis of learners' English. *Language Learning*, *61*, 647–672. http://doi-org.proxy1.lib.uwo.ca/10.1111/j.1467-9922.2010.00621.x

Li, J., & Schmitt, N. (2009). The acquisition of lexical phrases in academic writing: A longitudinal case study. *Journal of Second Language Writing*, *18*, 85–102. http://doi.org/10.1016/j.jslw.2009.02.001

Lindstromberg, S., Eyckmans, J., & Connabeer, R. (2016). A modified dictogloss for helping learners remember L2 academic English formulaic sequences for use in later writing. *English for specific purposes*, *41*, 12–21. http://doi.org/10.1016/j.esp.2015.08.002

Littlemore, J., Chen, P. T., Koester, A., & Barnden, J. (2011). Difficulties in metaphor comprehension faced by international students whose first language is not English. *Applied Linguistics*, *32*, 408–429. https://doi.org/10.1093/applin/amr009

Lomicka, L. (1998). To gloss or not to gloss: An investigation of reading comprehension online. *Language Learning & Technology*, *1*, 41–50.

Martinez, R., & Murphy, V. A. (2011). Effect of frequency and idiomaticity on second language reading comprehension. *TESOL Quarterly*, *45*, 267–290. https://doi.org/10.5054/tq.2011.247708

Mondria, J-A. (2003). The effects of inferring, verifying, and memorizing on the

retention of L2 word meanings: An experimental comparison of the "meaning-

inferred method" and the "meaning-given method." *Studies in Second*

Language Acquisition, *25*(4), 473–499.

https://doi.org/10.1017/S0272263103000202

Nagata, N. (1999). The effectiveness of computer-assisted interactive glosses.

Foreign Language Annals, *32*, 469–479.

https://doi.org/10.1111/j.1944-9720.1999.tb00876.x

Nassaji, H. (2003). L2 vocabulary learning from context: Strategies, knowledge

sources, and their relationship with success in L2 lexical inferencing. *TESOL*

Quarterly, *37*(4), 645–670. https://doi.org/10.2307/3588216

Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge

University Press.

Nesselhauf, N. (2003). The use of collocations by advanced learners of English and

some implications for teaching. *Applied Linguistics*, *24*, 223–242.

Nesselhauf, N. (2005). *Collocations in a learner corpus*. John Benjamins.

Paribakht, T. S., & Wesche, M. (1999). Reading and incidental vocabulary

acquisition: An introspective study of lexical inferencing. *Studies in Second*

Language Acquisition, *21*, 195–224.

Paivio, A. (1986). *Mental representations: A dual coding approach*. Oxford University

Press.

Parry, K. (1997). Vocabulary and comprehension: Two portraits. In J. Coady & T.

　　Huckin (Eds.), *Second language vocabulary acquisition: A rationale for*

　　pedagogy (pp. 54–68). Cambridge University Press.

Peters, E. (2012). Learning German formulaic sequences: The effect of two attention-

　　drawing techniques. *The Language Learning Journal*, *40*, 65–79.

　　https://doi.org/10.1080/09571736.2012.658224

Peters, E., Hulstijn, J. H., Sercu, L, & Lutjeharms, M. (2009). Learning L2 German

　　vocabulary through reading: The effect of three enhancement techniques

　　compared. *Language Learning*, *59*(1), 113–151.

　　https://doi.org/10.1111/j.1467-9922.2009.00502.x

Poole, R. (2012). Concordance-based glosses for academic vocabulary acquisition.

　　CALICO Journal, *29*, 679–693.

Pulido, D. (2004). The relationship between text comprehension and second language

　　incidental vocabulary acquisition: A matter of topic familiarity? *Language*

　　Learning, *54*, 469–523. https://doi.org/10.1111/j.0023-8333.2004.00263.x

Quené, H & van den Bergh, H. (2004). On multi-level modeling of data from repeated

　　measures designs: a tutorial. *Speech Communication*, *43*, 103–121.

　　https://doi.org/10.1016/j.specom.2004.02.004

Schmidt, R. W. (1995). Consciousness and foreign language learning: A tutorial on

　　the role of attention and awareness in learning. *Attention and Awareness in*

　　Foreign Language Learning, *9*, 1–63.

Schmidt, R. W. (2001). Attention. In P. Robinson (Ed.), *Cognition and second*

language instruction (pp. 3–32). Cambridge University Press.

Shin, D., & Nation, P. (2007). Beyond single words: The most frequent collocations

in spoken English. *ELT Journal*, *62*, 339–348.

http://dx.doi.org/10.1093/elt/ccm091

Siyanova-Chanturia, A. & Pellicer-Sanchez, A. (Eds.). (2018). *Understanding*

formulaic language: A second language acquisition perspective. Routledge.

http://doi-org.proxy1.lib.uwo.ca/10.4324/9781315206615

Siyanova-Chanturia, A., & Schmitt, N. (2008). L2 learner production and processing

of collocation: A multi-study perspective. *The Canadian Modern Language*

Review, *64*, 429–458.

The jamovi project (2021). *jamovi.* (Version 2.2) [Computer Software]. Retrieved

from https://www.jamovi.org.

Thomson, H., Boers, F., & Coxhead, A (2017). Replication research in pedagogical

approaches to spoken fluency and formulaic sequences: A call for replication

of Wood (2009) and Boers, Eychmans, Kappel, Stengers & Dececheleer

(2006). *Language Teaching* (only early view).

http://doi.org/10.1017/S0261444817000374

Tulving, E., & Kroll, N. (1995). Novelty assessment in the brain and long-term

memory encoding. *Psychonomic Bulletin and Review*, *2*, 387–390.

Wang, X., Boers, F., & Warren, P. (2020). Using literal underpinnings to help learners

remember figurative idioms: Does the connection need to be crystal-clear? In

A. M. Piquer-Píriz & R. Alejo-González (Eds.), *Metaphor in foreign language instruction* (221–239)*.* De Gruyter.

Warren, P, Boers, F., Grimshaw, G., & Siyanova-Chanturia, A. (2018). The effect of gloss type on learners' intake of new words during reading: Evidence from eye tracking. *Studies in Second Language Acquisition*, *40*, 883–906.

Watanabe, Y. (1997). Input, intake, and retention: Effects of increased processing on incidental learning of foreign language vocabulary. *Studies in Second Language Acquisition*, *19*, 287–307.

Webb, S. (2007). The effects of repetition on vocabulary knowledge. *Applied Linguistics*, *28*, 46–65.

Webb, S. (Ed.). (2020). *Approaches to learning, testing and researching L2 vocabulary*. John Benjamins Publishing Company.

Wood, D. (2010). *Formulaic language and second language speech fluency: Background, evidence and classroom applications*. Continuum.

Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge University Press.

Yanagisawa, A., Webb, S., & Uchihara T. (2020). How do different forms of glossing contribute to L2 vocabulary learning from reading? A meta-regression analysis. *Studies in Second Language Acquisition*, *42*, 411–438. https://doi.org/10.1017/S0272263119000688

Yoshii, M. (2006). L1 and L2 glosses: Their effects on incidental vocabulary learning. *Language Learning & Technology*, *10*, 85–101.

Yoshii, M., & Flaitz, J. (2002). Second language incidental vocabulary retention: The

effect of text and picture annotation types. *CALICO Journal*, *20*, 33–58.

Yoshii, M. (2013). Effects of gloss types on vocabulary learning through reading:

Comparison of single translation and multiple-choice gloss types. *CALICO*

Journal, *30*, 203–229.

Zhang, Z., & Webb, S. (2019). The effects of reading bilingual books on vocabulary

learning. *Reading in a Foreign Language*, *31*, 108–139.

Appendices

Appendix A

Reading Material 1 for Abstract-Only Group (Group A)

What Is the Nag Factor?

How parents respond to a child's pester power influences it.

By David J Bredehoft on October 15, 2021

When was the last time your children **nagged** you? Was it at the grocery store for candy? Or was it while you were driving, and they saw a McDonald's and would not stop pestering you to get a happy meal?

The Cambridge Dictionary defines pester power as, "The ability that children have to make their parents buy something, by asking for it many times until they get it." Pester power has also been referred to as "The Nag Factor."

Children use pester power to persuade parents to buy them things that they don't really need. The nag factor not only gets parents to buy things but is also used to manipulate parents to do things they normally would not. A 2002 survey found that most children repeatedly request an item until their parent **throws in the towel.** Some will ask more than 50 times for a particular product.

So, what can parents do in response to this? Behaviorist psychologist B. F. Skinner's basic conditioning principles may provide some answers for parents.

First: Almost everything is learned.
This includes children learning to nag and to pester parents to get what they want. In an environment where they see other children do it, they will **follow suit.**

Second: The most powerful tool in your behavioral toolbox is positive reinforcement.
Do you give in to your child's nagging? If so, you are reinforcing the very behavior you do not want.

It is understandable if you give in because no parent wants to be a **wet blanket**. Still, it is better to remove yourself and your child from the situation that triggered the

Nag: keep asking someone to do something.

Throw in the towel: (current) give up.

Follow suit: (current) doing the same as what someone else has just done.

Wet blanket: (current) someone who spoils other's excitement.

- 1 -

nagging and go to a neutral spot without such outside interference. This will **take the edge off** the situation, and once the child has calmed down, you should reward him or her with praise for calming down. Explain to him again what your previously established rule concerning nagging and whining is. Now you are reinforcing the appropriate behavior that you want to see happen. **Sticking to your guns** is extremely important.

Third: The second most powerful tool is ignoring inappropriate behavior.
Again, this is difficult to do but it is more powerful than most parents realize. Whenever possible, ignore your child's nagging. Go on about your business. Look the other way. Then, when your child is no longer nagging, give her positive verbal reinforcement for calming down, and for engaging in appropriate behavior.

Fourth: Punishment doesn't work very well.
Positive reinforcement is more effective than punishment. B. F. Skinner's research clearly points this out. In many cases, punishment backfires by drawing more attention to the very behavior you are trying to extinguish.

Fifth: Inappropriate behavior that is rewarded sometimes is the hardest to extinguish.
If you give in to the child's nagging sometimes while you ignore it other times, the nagging will continue. A useful analogy is the difference between **vending machines** and **slot machines**.
In the case of a vending machine, every time I put a coin in and punch the button, I expect to receive the item I ordered. If I do not, it means the machine does not work and I simply quit. In the case of a slot machine, however, I know I will often not receive anything in return for my coins, but there is an expectation that sometimes I will. So, I will keep putting in coins. Similarly, if your behavior is like the vending machine, the begging will stop much more quickly. If each time he begs, the same consequence occurs (leave the store immediately and take him home without any purchase), the begging will stop much more quickly.

Take the edge off:
(current) to reduce the intensity of an unpleasant situation.

Stick to your guns:
(current) refuse to change your decision.

Vending machine: you can buy drinks, snacks by putting coins into it; **Slot machine**: a machine used for gambling.

Appendix B

Reading Material 2 for Abstract-Only Group (Group A)

Earth Is on Fire

Our planet is burning, both literally and figuratively, because of climate change—and COVID is no excuse to ignore it.

I've never known an Earth that wasn't on fire.

I am 23 years old, and my entire generation has come of age in a world defined by climate change and other forms of environmental **degradation**. But, since the pandemic started, the fate of nature has become a side story to COVID.

degradation: getting damaged.

I know that saying this **goes against the grain**, but in the early days the pandemic seemed to be a positive thing for the future of our planet. While struggling to adjust to life in lockdowns, we were relieved to read that emissions levels had dipped, and that the skies over New Delhi and Los Angeles and Buenos Aires had cleared of smog. I smiled, as we all did, to notice that animals were traveling free through quiet, traffic-free cities. Nature seemed to be reclaiming spaces humans had abandoned. These story lines gave us faith in the planet's **resilience**.

go against the grain:
(current) difficult because it conflicts with established beliefs.

resilience: the ability to recover after difficulties.

Some even predicted that the experience would make us better managers of our world. Surely, seeing that a reduction in human economic activity was giving a chance to nature to recover would make us all appreciate the importance of reducing our harmful impact on the environment? Such predictions have turned out to be **wide of the mark**. We have somehow become *less* thoughtful in our daily choices—accepting that extra plastic bag at the supermarket, ordering takeout despite all the single-use containers and, if we are privileged enough, driving instead of taking public transport—because, well, "it is a global pandemic." Take a walk outside, and you will find masks and latex gloves littering our streets and beaches and parks, and they will eventually fill rivers, lakes and seas.

wide of the mark:
(current) very inaccurate.

It is as if the pandemic has suddenly given people everywhere even more of a license to dirty the world.

I fear that for every day it continues, today's young people will be paying the ecological price. I'm not just talking about those of us living in developed nations. I'm talking about children from impoverished families worldwide who are left **high and dry** because of unfair economic policies, climate change, and now also the coronavirus. Take for example, young indigenous people in Latin America, whose entire cultures (many of them predicated on harmony with nature) are being erased as their elders die of infection and as ranchers and miners violently and illegally drive them from their ancestral lands.

high and dry:
(current) in a difficult situation that you cannot do anything about

The rain forests in South America are disappearing at a record pace. There have been major oil spills in the Russian Arctic, Mauritius, Venezuela and California— terrible ecological disasters that are buried underneath headlines of COVID case numbers and mortality rates. It is as though the survival of our very planet has been **put on the back burner** because of the pandemic.

put on the back burner:
(current) something that does not need immediate attention.

It is hard for me to understand that some people can still so easily shrug off the realities of climate change and other obvious damage to the environment for which people are to blame, while for so many in my generation, it is a constant worry. We have set the planet on course for destruction, and this is no **dummy run**.

dummy run:
(current) a test.

However, I like to think that, perhaps if we all live a little lighter, if we listen to those who are in harmony with the land and if we seek **solace** in all that there is to love in the world, nature might meet us halfway. The planet and our fates hang in the balance.

solace: comfort, hope, consolation.

Appendix C

Reading Material 1 for Literal-Only Group (Group B)

What Is the Nag Factor?

How parents respond to a child's pester power influences it.

By David J Bredehoft on October 15, 2021

When was the last time your children **nagged** you? Was it at the grocery store for candy? Or was it while you were driving, and they saw a McDonald's and would not stop pestering you to get a happy meal?

Nag：keep asking someone to do something.

The Cambridge Dictionary defines pester power as, "The ability that children have to make their parents buy something, by asking for it many times until they get it." Pester power has also been referred to as "The Nag Factor."

Children use pester power to persuade parents to buy them things that they don't really need. The nag factor not only gets parents to buy things but is also used to manipulate parents to do things they normally would not. A 2002 survey found that most children repeatedly request an item until their parent **throws in the towel.** Some will ask more than 50 times for a particular product.

Throw in the towel: (origin) in a boxing match, if a coach notices his fighter is losing badly, he may throw a towel into to the ring to signal surrender.

So, what can parents do in response to this? Behaviorist psychologist B. F. Skinner's basic conditioning principles may provide some answers for parents.

First: Almost everything is learned.
This includes children learning to nag and to pester parents to get what they want. In an environment where they see other children do it, they will **follow suit**.

Follow suit: (origin) in a card game, you play a card of the same kind as the previous player.

Second: The most powerful tool in your behavioral toolbox is positive reinforcement.
Do you give in to your child's nagging? If so, you are reinforcing the very behavior you do not want.

It is understandable if you give in because no parent wants to be a **wet blanket**. Still, it is better to remove yourself and your child from the situation that triggered the

Wet blanket: (origin) used to put out a campfire.

nagging and go to a neutral spot without such outside interference. This will **take the edge off** the situation, and once the child has calmed down, you should reward him or her with praise for calming down. Explain to him again what your previously established rule concerning nagging and whining is. Now you are reinforcing the appropriate behavior that you want to see happen. **Sticking to your guns** is extremely important.

Third: The second most powerful tool is ignoring inappropriate behavior.
Again, this is difficult to do but it is more powerful than most parents realize. Whenever possible, ignore your child's nagging. Go on about your business. Look the other way. Then, when your child is no longer nagging, give her positive verbal reinforcement for calming down, and for engaging in appropriate behavior.

Fourth: Punishment doesn't work very well.
Positive reinforcement is more effective than punishment. B. F. Skinner's research clearly points this out. In many cases, punishment backfires by drawing more attention to the very behavior you are trying to extinguish.

Fifth: Inappropriate behavior that is rewarded sometimes is the hardest to extinguish.
If you give in to the child's nagging sometimes while you ignore it other times, the nagging will continue. A useful analogy is the difference between **vending machines** and **slot machines**.
In the case of a vending machine, every time I put a coin in and punch the button, I expect to receive the item I ordered. If I do not, it means the machine does not work and I simply quit. In the case of a slot machine, however, I know I will often not receive anything in return for my coins, but there is an expectation that sometimes I will. So, I will keep putting in coins. Similarly, if your behavior is like the vending machine, the begging will stop much more quickly. If each time he begs, the same consequence occurs (leave the store immediately and take him home without any purchase), the begging will stop much more quickly.

Take the edge off: (origin) make a knife or sword less sharp.

Stick to your guns: (origin) on the battlefield, soldiers should stay with their cannons (guns), even if they are under attack.

Vending machine: you can buy drinks, snacks by putting coins into it; **Slot machine**: a machine used for gambling.

Appendix D

Reading Material 2 for Literal-Only Group (Group B)

Earth Is on Fire

Our planet is burning, both literally and figuratively, because of climate change—and COVID is no excuse to ignore it.

I've never known an Earth that wasn't on fire.

I am 23 years old, and my entire generation has come of age in a world defined by climate change and other forms of environmental **degradation**. But, since the pandemic started, the fate of nature has become a side story to COVID.

degradation: getting damaged.

I know that saying this **goes against the grain**, but in the early days the pandemic seemed to be a positive thing for the future of our planet. While struggling to adjust to life in lockdowns, we were relieved to read that emissions levels had dipped, and that the skies over New Delhi and Los Angeles and Buenos Aires had cleared of smog. I smiled, as we all did, to notice that animals were traveling free through quiet, traffic-free cities. Nature seemed to be reclaiming spaces humans had abandoned. These story lines gave us faith in the planet's **resilience**.

go against the grain: (origin) cutting a piece of wood is harder when you try to do it against its grain (the direction of its fibers)

resilience: the ability to recover after difficulties.

Some even predicted that the experience would make us better managers of our world. Surely, seeing that a reduction in human economic activity was giving a chance to nature to recover would make us all appreciate the importance of reducing our harmful impact on the environment? Such predictions have turned out to be **wide of the mark**. We have somehow become *less* thoughtful in our daily choices—accepting that extra plastic bag at the supermarket, ordering takeout despite all the single-use containers and, if we are privileged enough, driving instead of taking public transport—because, well, "it is a global pandemic." Take a walk outside, and you will find masks and latex gloves littering our streets and beaches and parks, and they will eventually fill rivers, lakes and seas.

wide of the mark: (origin) the mark is the target you aim at in shooting; if you miss badly, your arrow will fall at a wide distance from it

It is as if the pandemic has suddenly given people everywhere even more of a license to dirty the world.

I fear that for every day it continues, today's young people will be paying the ecological price. I'm not just talking about those of us living in developed nations. I'm talking about children from impoverished families worldwide who are left **high and dry** because of unfair economic policies, climate change, and now also the coronavirus. Take for example, young indigenous people in Latin America, whose entire cultures (many of them predicated on harmony with nature) are being erased as their elders die of infection and as ranchers and miners violently and illegally drive them from their ancestral lands.

high and dry: (origin) a boat is stuck on land or on a sandbank.

The rain forests in South America are disappearing at a record pace. There have been major oil spills in the Russian Arctic, Mauritius, Venezuela and California—terrible ecological disasters that are buried underneath headlines of COVID case numbers and mortality rates. It is as though the survival of our very planet has been **put on the back burner** because of the pandemic.

put on the back burner: (origin) when we cook, we put things to simmer at the back of the stove, so we can use the front burners to attend to other pots or pans.

It is hard for me to understand that some people can still so easily shrug off the realities of climate change and other obvious damage to the environment for which people are to blame, while for so many in my generation, it is a constant worry. We have set the planet on course for destruction, and this is no **dummy run**.

dummy run: (origin) in a car crash test, a dummy is used instead of a real person.

However, I like to think that, perhaps if we all live a little lighter, if we listen to those who are in harmony with the land and if we seek **solace** in all that there is to love in the world, nature might meet us halfway. The planet and our fates hang in the balance.

solace: comfort, hope, consolation.

Appendix E

Reading Material 1 for Abstract-Plus-Literal Group (Group C)

What Is the Nag Factor?

How parents respond to a child's pester power influences it.

By David J Bredehoft on October 15, 2021

When was the last time your children **nagged** you? Was it at the grocery store for candy? Or was it while you were driving, and they saw a McDonald's and would not stop pestering you to get a happy meal?

| **Nag**：keep asking someone to do something. |

The Cambridge Dictionary defines pester power as, "The ability that children have to make their parents buy something, by asking for it many times until they get it." Pester power has also been referred to as "The Nag Factor."

Children use pester power to persuade parents to buy them things that they don't really need. The nag factor not only gets parents to buy things but is also used to manipulate parents to do things they normally would not. A 2002 survey found that most children repeatedly request an item until their parent **throws in the towel.** Some will ask more than 50 times for a particular product.

Throw in the towel: (current) give up; (origin) in a boxing match, if a coach notices his fighter is losing badly, he may throw a towel into to the ring to signal surrender.

So, what can parents do in response to this? Behaviorist psychologist B. F. Skinner's basic conditioning principles may provide some answers for parents.

First: Almost everything is learned.
This includes children learning to nag and to pester parents to get what they want. In an environment where they see other children do it, they will **follow suit**.

Follow suit: (current) doing the same as what someone else has just done; (origin) in a card game, you play a card of the same kind as the previous player.

Second: The most powerful tool in your behavioral toolbox is positive reinforcement.
Do you give in to your child's nagging? If so, you are reinforcing the very behavior you do not want.

It is understandable if you give in because no parent wants to be a **wet blanket**. Still, it is better to remove yourself and your child from the situation that triggered the

Wet blanket: (current) someone who spoils others' excitement; (origin) used to put out a campfire.

nagging and go to a neutral spot without such outside interference. This will **take the edge off** the situation, and once the child has calmed down, you should reward him or her with praise for calming down. Explain to him again what your previously established rule concerning nagging and whining is. Now you are reinforcing the appropriate behavior that you want to see happen. **Sticking to your guns** is extremely important.

Third: The second most powerful tool is ignoring inappropriate behavior.
Again, this is difficult to do but it is more powerful than most parents realize. Whenever possible, ignore your child's nagging. Go on about your business. Look the other way. Then, when your child is no longer nagging, give her positive verbal reinforcement for calming down, and for engaging in appropriate behavior.

Fourth: Punishment doesn't work very well.
Positive reinforcement is more effective than punishment. B. F. Skinner's research clearly points this out. In many cases, punishment backfires by drawing more attention to the very behavior you are trying to extinguish.

Fifth: Inappropriate behavior that is rewarded sometimes is the hardest to extinguish.
If you give in to the child's nagging sometimes while you ignore it other times, the nagging will continue. A useful analogy is the difference between **vending machines** and **slot machines**.
In the case of a vending machine, every time I put a coin in and punch the button, I expect to receive the item I ordered. If I do not, it means the machine does not work and I simply quit. In the case of a slot machine, however, I know I will often not receive anything in return for my coins, but there is an expectation that sometimes I will. So, I will keep putting in coins. Similarly, if your behavior is like the vending machine, the begging will stop much more quickly. If each time he begs, the same consequence occurs (leave the store immediately and take him home without any purchase), the begging will stop much more quickly.

Take the edge off: (current) to reduce the intensity of an unpleasant situation; (origin) make a knife or sword less sharp.

Stick to your guns: (current) refuse to change your decision; (origin) on the battlefield, soldiers should stay with their cannons (guns), even if they are under attack.

Vending machine: you can buy drinks, snacks by putting coins into it; **Slot machine**: a machine used for gambling.

Appendix F

Reading Material 2 for Abstract-Plus-Literal Group (Group C)

Earth Is on Fire

Our planet is burning, both literally and figuratively, because of climate change—and COVID is no excuse to ignore it.

I've never known an Earth that wasn't on fire.

I am 23 years old, and my entire generation has come of age in a world defined by climate change and other forms of environmental **degradation**. But, since the pandemic started, the fate of nature has become a side story to COVID.

degradation: getting damaged.

I know that saying this **goes against the grain**, but in the early days the pandemic seemed to be a positive thing for the future of our planet. While struggling to adjust to life in lockdowns, we were relieved to read that emissions levels had dipped, and that the skies over New Delhi and Los Angeles and Buenos Aires had cleared of smog. I smiled, as we all did, to notice that animals were traveling free through quiet, traffic-free cities. Nature seemed to be reclaiming spaces humans had abandoned. These story lines gave us faith in the planet's **resilience**.

go against the grain: (current) difficult because it conflicts with established beliefs; (origin) cutting a piece of wood is harder when you try to do it against its grain (the direction of its fibers).

resilience: the ability to recover after difficulties.

Some even predicted that the experience would make us better managers of our world. Surely, seeing that a reduction in human economic activity was giving a chance to nature to recover would make us all appreciate the importance of reducing our harmful impact on the environment? Such predictions have turned out to be **wide of the mark**. We have somehow become *less* thoughtful in our daily choices—accepting that extra plastic bag at the supermarket, ordering takeout despite all the single-use containers and, if we are privileged enough, driving instead of taking public transport—because, well, "it is a global pandemic." Take a walk outside, and you will find masks and latex gloves littering our streets and beaches and parks, and they will eventually fill rivers, lakes and seas.

wide of the mark: (current) very inaccurate; (origin) the mark is the target you aim at in shooting; if you miss badly, your arrow will fall at a wide distance from it.

It is as if the pandemic has suddenly given people everywhere even more of a license to dirty the world.

I fear that for every day it continues, today's young people will be paying the ecological price. I'm not just talking about those of us living in developed nations. I'm talking about children from impoverished families worldwide who are left **high and dry** because of unfair economic policies, climate change, and now also the coronavirus. Take for example, young indigenous people in Latin America, whose entire cultures (many of them predicated on harmony with nature) are being erased as their elders die of infection and as ranchers and miners violently and illegally drive them from their ancestral lands.

high and dry: (current) in a difficult situation that you cannot do anything about; (origin) a boat is stuck on land or on a sandbank.

The rain forests in South America are disappearing at a record pace. There have been major oil spills in the Russian Arctic, Mauritius, Venezuela and California— terrible ecological disasters that are buried underneath headlines of COVID case numbers and mortality rates. It is as though the survival of our very planet has been **put on the back burner** because of the pandemic.

put on the back burner: (current) something that does not need immediate attention; (origin) when we cook, we put things to simmer at the back of the stove, so we can use the front burners to attend to other pots or pans.

It is hard for me to understand that some people can still so easily shrug off the realities of climate change and other obvious damage to the environment for which people are to blame, while for so many in my generation, it is a constant worry. We have set the planet on course for destruction, and this is no **dummy run**.

dummy run: (current) a test; (origin) in a car crash test, a dummy is used instead of a real person.

However, I like to think that, perhaps if we all live a little lighter, if we listen to those who are in harmony with the land and if we seek **solace** in all that there is to love in the world, nature might meet us halfway. The planet and our fates hang in the balance.

solace: comfort, hope, consolation.

Appendix G

Content-Related Questions After Reading Material 1

Content-related Questions

In this article, five basic conditional principles are provided to parents. Please read it again and select the response that best indicates your opinion about the statements below.

1. *Almost everything is learned.*

	1	2	3	4	5	6	
Totally Disagree	O	O	O	O	O	O	Totally Agree

2. *The most powerful tool in your behavioral toolbox is positive reinforcement.*

	1	2	3	4	5	6	
Totally Disagree	O	O	O	O	O	O	Totally Agree

3. *The second most powerful tool is ignoring inappropriate behavior.*

	1	2	3	4	5	6	
Totally Disagree	O	O	O	O	O	O	Totally Agree

4. *Punishment doesn't work very well.*

	1	2	3	4	5	6	
Totally Disagree	O	O	O	O	O	O	Totally Agree

5. *Inappropriate behavior that is rewarded sometimes is the hardest to extinguish.*

	1	2	3	4	5	6	
Totally Disagree	O	O	O	O	O	O	Totally Agree

Appendix H

Content-Related Questions After Reading Material 2

Content-related Questions

Please read the following statements again and select the response that best indicates your attitude.

1. In the early days the pandemic seemed to be a positive thing for the future of our planet.

	1	2	3	4	5	6	
Totally Disagree	O	O	O	O	O	O	Totally Agree

2. During this pandemic time, we have somehow become *less* thoughtful in our daily choices—accepting that extra plastic bag at the supermarket…

	1	2	3	4	5	6	
Totally Disagree	O	O	O	O	O	O	Totally Agree

3. It is as if the pandemic has suddenly given people everywhere even more of a license to dirty the world.

	1	2	3	4	5	6	
Totally Disagree	O	O	O	O	O	O	Totally Agree

4. If we all live a little lighter, if we listen to those who are in harmony with the land and if we seek solace in all that there is to love in the world, nature might meet us halfway.

	1	2	3	4	5	6	
Totally Disagree	O	O	O	O	O	O	Totally Agree

Appendix I

Immediate Post-test for Reading Material 1

Gap-fills

Complete the sentence by filling in the blanks. Each blank stands for one word.

When was the last time your children _____you? Was it at the grocery store for candy? Or was it while you were driving, and they saw a McDonald's and would not stop pestering you to get a happy meal?

The Cambridge Dictionary defines pester power as, "The ability that children have to make their parents buy something, by asking for it many times until they get it." Pester power has also been referred to as "The Nag Factor."

Children use pester power to persuade parents to buy them things that they don't really need. The nag factor not only gets parents to buy things but is also used to manipulate parents to do things they normally would not. A 2002 survey found that most children repeatedly request an item until their parent **throws in the** _____. Some will ask more than 50 times for a particular product.

So, what can parents do in response to this? Behaviorist psychologist B. F. Skinner's basic conditioning principles may provide some answers for parents.

First: Almost everything is learned.
This includes children learning to nag and to pester parents to get what they want. In an environment where they see other children do it, they will **follow** _____.

Second: The most powerful tool in your behavioral toolbox is positive reinforcement.
Do you give in to your child's nagging? If so, you are reinforcing the very behavior you do not want.

It is understandable if you give in because no parent wants to be a **wet** _____. Still, it is better to remove yourself and your child from the situation that triggered the nagging and go to a neutral spot without such outside interference. This will **take the** _____ **off** the situation, and once the child has calmed down, you should reward him or her with praise for calming down. Explain to him again what your previously established rule concerning nagging and whining is. Now you are reinforcing the appropriate behavior that you want to see happen. **Sticking to your** _____ is extremely important.

Third: The second most powerful tool is ignoring inappropriate behavior.
Again, this is difficult to do but it is more powerful than most parents realize. Whenever possible, ignore your child's nagging. Go on about your business. Look the other way. Then, when your child is no longer nagging, give her positive verbal reinforcement for calming down, and for engaging in appropriate behavior.

Fourth: Punishment doesn't work very well.
Positive reinforcement is more effective than punishment. B. F. Skinner's research
clearly points this out. In many cases, punishment backfires by drawing
more attention to the very behavior you are trying to extinguish.

Fifth: Inappropriate behavior that is rewarded sometimes is the hardest to extinguish.
If you give in to the child's nagging sometimes while you ignore it other times, the
nagging will continue. A useful analogy is the difference between _____
machines and _____ **machines**.

In the case of a vending machine, every time I put a coin in and punch the button, I
expect to receive the item I ordered. If I do not, it means the machine does not work
and I simply quit. In the case of a slot machine, however, I know I will often not
receive anything in return for my coins, but there is an expectation that sometimes I
will. So, I will keep putting in coins. Similarly, if your behavior is like the vending
machine, the begging will stop much more quickly. If each time he begs, the same
consequence occurs (leave the store immediately and take him home without any
purchase), the begging will stop much more quickly.

Meaning Explanations

Please read the sentences and **explain the meaning of the word or phrases in bold**.

1. When was the last time your children **nagged** you? Was it at the grocery store for candy? Or was it while you were driving, and they saw a McDonald's and would not stop pestering you to get a happy meal?

   _____

   _____

   Did you know this word before?

   ☐ Yes

   ☐ No

2. A 2002 survey found that most children repeatedly request an item until their parent **throws in the towel**.

   _____

   _____

   Did you know this phrase before?

   ☐ Yes

   ☐ No

3. This includes children learning to nag and to pester parents to get what they want. In an environment where they see other children do it, they will **follow suit**.

   _____

   _____

   Did you know this phrase before?

   ☐ Yes

   ☐ No

4. Do you give in to your child's nagging? If so, you are reinforcing the very behavior you do not want. It is understandable if you give in because no parent wants to be a **wet blanket**.

   _____

   _____

Did you know this phrase before?

☐   Yes

☐   No

5. Still, it is better to remove yourself and your child from the situation that triggered the nagging and go to a neutral spot without such outside interference. This will **take the edge off** the situation, and once the child has calmed down, you should reward him or her with praise for calming down.

_____

_____

Did you know this phrase before?

☐   Yes

☐   No

6. Explain to him again what your previously established rule concerning nagging and whining is. Now you are reinforcing the appropriate behavior that you want to see happen. **Sticking to your guns** is extremely important.

_____

_____

Did you know this phrase before?

☐   Yes
☐   No

7. If you give in to the child's nagging sometimes while you ignore it other times, the nagging will continue. A useful analogy is the difference between **vending machines** and **slot machines**.

_____

_____

_____

Did you know this phrase "vending machine" before?

☐   Yes
☐   No

Did you know this phrase "slot machine" before?

☐   Yes

☐   No

Appendix J

Immediate Post-test for Reading Material 2

Gap-fills

Complete the sentences by filling in the blanks. Each blank stands for one word.

I've never known an Earth that wasn't on fire.

I am 23 years old, and my entire generation has come of age in a world defined by climate change and other forms of environmental _____. But, since the pandemic started, the fate of nature has become a side story to COVID.

I know that saying this **goes against the** _____, but in the early days the pandemic seemed to be a positive thing for the future of our planet. While struggling to adjust to life in lockdowns, we were relieved to read that emissions levels had dipped, and that the skies over New Delhi and Los Angeles and Buenos Aires had cleared of smog. I smiled, as we all did, to notice that animals were traveling free through quiet, traffic-free cities. Nature seemed to be reclaiming spaces humans had abandoned. These story lines gave us faith in the planet's _____.

Some even predicted that the experience would make us better managers of our world. Surely, seeing that a reduction in human economic activity was giving a chance to nature to recover would make us all appreciate the importance of reducing our harmful impact on the environment? Such predictions have turned out to be **wide of the** _____. We have somehow become *less* thoughtful in our daily choices—accepting that extra plastic bag at the supermarket, ordering takeout despite all the single-use containers and, if we are privileged enough, driving instead of taking public transport—because, well, "it is a global pandemic." Take a walk outside, and you will find masks and latex gloves littering our streets and beaches and parks, and they will eventually fill rivers, lakes and seas.

It is as if the pandemic has suddenly given people everywhere even more of a license to dirty the world.
I fear that for every day it continues, today's young people will be paying the ecological price. I'm not just talking about those of us living in developed nations. I'm talking about children from impoverished families worldwide who are left **high and** _____ because of unfair economic policies, climate change, and now also the coronavirus. Take for example, young indigenous people in Latin America, whose entire cultures (many of them predicated on harmony with nature) are being erased as their elders die of infection and as ranchers and miners violently and illegally drive them from their ancestral lands.

The rain forests in South America are disappearing at a record pace. There have been major oil spills in the Russian Arctic, Mauritius, Venezuela and California—terrible ecological disasters that are buried underneath headlines of COVID case numbers and mortality rates. It is as though the survival of our very planet has been **put on the back** _____ because of the pandemic.

It is hard for me to understand that some people can still so easily shrug off the realities of climate change and other obvious damage to the environment for which people are to blame, while for so many in my generation, it is a constant worry. We have set the planet on course for destruction, and this is no **dummy** _____.

However, I like to think that, perhaps if we all live a little lighter, if we listen to those who are in harmony with the land and if we seek _____ in all that there is to love in the world, nature might meet us halfway. The planet and our fates hang in the balance.

Meaning Explanations

Please read the sentences again and **explain the meaning of the words or phrases in bold**.

1. I am 23 years old, and my entire generation has come of age in a world defined by climate change and other forms of environmental **degradation**.

   _____

   _____

   Did you know this word before?

   ☐   Yes
   ☐   No

2. I know that saying this **goes against the grain**, but in the early days the pandemic seemed to be a positive thing for the future of our planet.

   _____

   _____

   Did you know this phrase before?

   ☐   Yes
   ☐   No

3. Nature seemed to be reclaiming spaces humans had abandoned. These story lines gave us faith in the planet's **resilience**.

   _____

   _____

   Did you know this word before?

   ☐   Yes
   ☐   No

4. Some even predicted that the experience would make us better managers of our world … Such predictions have turned out to be **wide of the mark**.

   _____

   _____

   Did you know this phrase before?

   ☐   Yes
   ☐   No

5. I'm talking about children from impoverished families worldwide who are left **high and dry** because of unfair economic policies, climate change, and now also the coronavirus.

_____

_____

Did you know this phrase before?

☐ Yes

☐ No

6. Terrible ecological disasters have been buried underneath headlines of COVID case numbers and mortality rates. It is as though the survival of our very planet has been **put on the back burner** because of the pandemic.

_____

_____

Did you know this phrase before?

☐ Yes

☐ No

7. We have set the planet on course for destruction, and this is no **dummy run**.

_____

_____

Did you know this phrase before?

☐ Yes

☐ No

8. However, I like to think that, perhaps if we all live a little lighter, if we listen to those who are in harmony with the land and if we seek **solace** in all that there is to love in the world, nature might meet us halfway.

_____

_____

Did you know this word before?

☐ Yes

☐ No

Appendix K

Delayed Post-test for Reading Material 1

Meaning Explanations

Please write down the meaning of the following word/phrases.

1. Throw in the towel

    _____

    _____

2. Follow suit

    _____

    _____

3. A wet blanket

    _____

    _____

4. Take the edge off something

    _____

    _____

5. Stick to your guns

    _____

    _____

Multiple-choice Test

Please choose the most suitable meaning for the following words/phrases.

1. Throw in the towel
    - ☐ Refuse someone's request
    - ☐ Give up doing something
    - ☐ Postpone something to deal with something else
    - ☐ Make someone angry

2. Follow suit
    - ☐ Recover from damage
    - ☐ Obey people with authority
    - ☐ Do the same as what someone else has just done
    - ☐ Follow a particular dressing code

3. A wet blanket
    - ☐ Someone who spoils others' excitement
    - ☐ Someone who is busy all day but achieves nothing
    - ☐ Something that is heavier than expected
    - ☐ The ability to recover after difficulties

4. Take the edge off something
    - ☐ Create more freedom of expression
    - ☐ Say something that conflicts with established beliefs
    - ☐ Be rescued from a dangerous situation
    - ☐ Reduce the intensity of an unpleasant situation

5. Stick to your guns
    - ☐ Hold your weapon firmly so you're ready to use it
    - ☐ Refuse to change your position or decision
    - ☐ Make a very inaccurate prediction
    - ☐ Wait eagerly for a chance to take revenge

Appendix L

Delayed Post-test for Reading Material 2

Meaning Explanations

Please write down the meaning of the following phrases.

1. Go against the grain

   _____

   _____

2. Wide of mark

   _____

   _____

3. High and dry

   _____

   _____

4. Put on the back burner

   _____

   _____

5. Dummy run

   _____

   _____

Multiple-choice Test

Please choose the most suitable meaning for the following phrases.

1. Go against the grain
   - ☐ Be in a difficult situation
   - ☐ Damage the harvest
   - ☐ Conflict with established beliefs
   - ☐ Give up doing something

2. Wide of the mark
   - ☐ Very ambitious
   - ☐ Very distracting
   - ☐ Very delayed
   - ☐ Very inaccurate

3. High and dry
   - ☐ In a period of recovery
   - ☐ In a difficult situation
   - ☐ In a position of power
   - ☐ In high altitude areas

4. Put something on the back burner
   - ☐ Stop doing something harmful before it is too late
   - ☐ Postpone something to give priority to something else
   - ☐ Make a conflict or argument less intense
   - ☐ Give something time to recover from damage

5. A dummy run
   - ☐ A test
   - ☐ A disappointment
   - ☐ An escape
   - ☐ A bad mistake

Appendix M

Interview Guide Questions for Abstract-Only Group

Thank you again for your participation in this research. I would like to ask you some questions related to the texts you read as part of this study. Please note that our conversation will be recorded but that all the information collected from you will be kept confidential. Do you have any questions? If not, we will begin our interview.

1. I will be showing you the texts you read again, but before doing so I'd like to ask you if you can remember any supporting information that was added to the text to help you understand vocabulary. If so, can you describe what this support looked like?

2. Now I will show you the texts again.

   Figurative-meaning glosses:

   (a) Did you read the glosses?

   (b) Which expressions did you not yet know the meaning of?

   (c) Were the glosses helpful, or would you have been able to guess the meanings anyhow? How?

   (d) The expressions are idioms, that is, figurative phrases that were once used literally. Can you guess their original, literal use?

3. I'm now going to show you the text versions that were used in the other participant groups. As you can tell, the support they give for comprehension of idiomatic expressions varies. Using your own intuition and experience as language learners/ teachers, which of them do you think will lead to the best retention of the expressions in memory? Try to rank the 4 text versions from most effective to least effective for remembering the expressions. You can think aloud as you do this. As you can now guess, the, relative effectiveness of different gloss types is what we wished to test in this study. We'll offer to share the findings with you and the other participants as soon as we have completed the data analysis.

1   2   3   4

4. I did not tell you in advance that a test about the idiomatic expressions would follow the reading activities. This was because the study was meant to be about incidental learning (something you will recognize from the course about vocabulary teaching and learning that you took in the fall term). Did you nonetheless perhaps expect such a test? Did you make an effort to commit the expressions to memory, just in case? After being tested straight after the reading activities, did you suspect a delayed test would follow?

5. Language courses sometimes include texts accompanied by glosses to clarify vocabulary. Is this something you remember from your own language learning or teaching experience? Do you find reading with glosses useful for language learning? Why (not)? Can you remember reading glosses of the kind we used in the experiment, notably glosses that inform students about the origin of expressions instead of simply clarifying the current meaning?

6. Are there any other thoughts you would like to share about the study and how you experienced the activities and tests?

Thank you so much for your help!

Appendix N

Interview Guide Questions for Literal-Only Group

Thank you again for your participation in this research. I would like to ask you some questions related to the texts you read as part of this study. Please note that our conversation will be recorded but that all the information collected from you will be kept confidential. Do you have any questions? If not, we will begin our interview.

1. I will be showing you the texts you read again, but before doing so I'd like to ask you if you can remember any supporting information that was added to the text to help you understand vocabulary. If so, can you describe what this support looked like?

2. Now I will show you the texts again.

   Literal-meaning glosses:

   (a) Did you read the glosses?

   (b) Which expressions did you not yet know the meaning of?

   (c) Were the glosses helpful? Did the information help you understand the expressions in the text?

   (d) You probably noticed that the information in the gloss did not directly clarify the actual meaning of the expressions in the text, but just their origin. Did you find it hard to make the connection? Did it make you feel curious about the current meaning of the expressions? Do you find information about the origin of the expressions useful? Why (not)?

3. I'm now going to show you the text versions that were used in the other participant groups. As you can tell, the support they give for comprehension of idiomatic expressions varies. Using your own intuition and experience as language learners/ teachers, which of them do you think will lead to the best retention of the expressions in memory? Try to rank the 4 text versions from most effective to least effective for remembering the expressions. You can think aloud as you do this. As you can now guess, the, relative effectiveness of different gloss types is what we wished to test in this study. We'll offer to share the findings with you and the other participants as soon as we have completed the data analysis.

    1   2   3   4

4. I did not tell you in advance that a test about the idiomatic expressions would follow the reading activities. This was because the study was meant to be about incidental learning (something you will recognize from the course about vocabulary teaching and learning that you took in the fall term). Did you nonetheless perhaps expect such a test? Did you make an effort to commit the expressions to memory, just in case? After being tested straight after the reading activities, did you suspect a delayed test would follow?

5. Language courses sometimes include texts accompanied by glosses to clarify vocabulary. Is this something you remember from your own language learning or teaching experience? Do you find reading with glosses useful for language learning? Why (not)? Can you remember reading glosses of the kind we used in the experiment, notably glosses that inform students about the origin of expressions instead of simply clarifying the current meaning?

6. Are there any other thoughts you would like to share about the study and how you experienced the activities and tests?

Thank you so much for your help!

Appendix O

Interview Guide Questions for Abstract-Plus-Literal Group

Thank you again for your participation in this research. I would like to ask you some questions related to the texts you read as part of this study. Please note that our conversation will be recorded but that all the information collected from you will be kept confidential. Do you have any questions? If not, we will begin our interview.

1. I will be showing you the texts you read again, but before doing so I'd like to ask you if you can remember any supporting information that was added to the text to help you understand vocabulary. If so, can you describe what this support looked like?

2. Now I will show you the texts again.

   Figurative + literal glosses:

   (a) Did you read the glosses? Most had two parts, one explaining the actual meaning of the expression and another part giving background information about its origin. Did you read both?

   (b) Which expressions did you not yet know the meaning of?

   (c) Were the glosses helpful, or would you have been able to guess the meanings anyhow? How?

   (d) Returning to the information in the glosses about the origin of the expressions, do you find this useful information? Why (not)?

3. I'm now going to show you the text versions that were used in the other participant groups. As you can tell, the support they give for comprehension of idiomatic expressions varies. Using your own intuition and experience as language learners/ teachers, which of them do you think will lead to the best retention of the expressions in memory? Try to rank the 4 text versions from most effective to least effective for remembering the expressions. You can think aloud as you do this. As you can now guess, the, relative effectiveness of different gloss types is what we wished to test in this study. We'll offer to share the findings with you and the other participants as soon as we have completed the data analysis.

1   2   3   4

4. I did not tell you in advance that a test about the idiomatic expressions would follow the reading activities. This was because the study was meant to be about incidental learning (something you will recognize from the course about vocabulary teaching and learning that you took in the fall term). Did you nonetheless perhaps expect such a test? Did you make an effort to commit the expressions to memory, just in case? After being tested straight after the reading activities, did you suspect a delayed test would follow?

5. Language courses sometimes include texts accompanied by glosses to clarify vocabulary. Is this something you remember from your own language learning or teaching experience? Do you find reading with glosses useful for language learning? Why (not)? Can you remember reading glosses of the kind we used in the experiment, notably glosses that inform students about the origin of expressions instead of simply clarifying the current meaning?

6. Are there any other thoughts you would like to share about the study and how you experienced the activities and tests?

Thank you so much for your help!

Appendix P

Ethics Approval Notice

Western Research

Date: 25 January 2022

To: Professor Frank Boers

Project ID: 120183

Study Title: Reading idiomatic discourse in a second language with different types of support

Short Title: L2 Reading with different types of support

Application Type: NMREB Initial Application

Review Type: Delegated

Full Board Reporting Date: February 4 2022

Date Approval Issued: 25/Jan/2022 23:13

REB Approval Expiry Date: 25/Jan/2023

Dear Professor Frank Boers

The Western University Non-Medical Research Ethics Board (NMREB) has reviewed and approved the WREM application form for the above mentioned study, as of the date noted above. NMREB approval for this study remains valid until the expiry date noted above, conditional to timely submission and acceptance of NMREB Continuing Ethics Review.

This research study is to be conducted by the investigator noted above. **All other required institutional approvals and mandated training must also be obtained prior to the conduct of the study.**

Documents Approved:

Document Name	Document Type	Document Date	Document Version
Reading Material 1 with Literal Gloss of Idioms	Other Data Collection Instruments		
Reading Material 1 with Abstract Gloss of Idioms	Other Data Collection Instruments		
Reading Material 1 with Both Literal Gloss and Abstract Gloss of Idioms	Other Data Collection Instruments		
Reading Material 1 with No Gloss of Idioms	Other Data Collection Instruments		
Reading Material 2 with Literal Gloss of Idioms	Other Data Collection Instruments		
Reading Material 2 with Abstract Gloss of Idioms	Other Data Collection Instruments		
Reading Material 2 with No Gloss of Idioms	Other Data Collection Instruments		
Reading Material 2 with Both Literal Gloss and Abstract Gloss of Idioms	Other Data Collection Instruments		
Concept-related Questions for Reading Material 1	Other Data Collection Instruments		
Concept-related Questions for Reading Material 2	Other Data Collection Instruments		
Immediate Post-test for Reading Material 1	Other Data Collection Instruments		
Immediate Post-test for Reading Material 2	Other Data Collection Instruments		

Delayed Post-test for Reading Material 1	Other Data Collection Instruments		
Delayed Post-test for Reading Material 2	Other Data Collection Instruments		
DEBRIEFING FORM	Debriefing document		
Interview Guide 23 Dec	Other Data Collection Instruments	23/Dec/2021	1
verbal recruitment script_revised	Oral Script	24/Dec/2021	2
Letter of Information and Consent Forms (LOI_C)_revised	Written Consent/Assent	24/Dec/2021	2
Interview Guide 23 Dec	Interview Guide	23/Dec/2021	1

Documents Acknowledged:

Document Name	Document Type	Document Date	Document Version
Outline of procedures	Supplementary Tables/Figures		

No deviations from, or changes to the protocol should be initiated without prior written approval from the NMREB, except when necessary to eliminate immediate hazard(s) to study participants or when the change(s) involves only administrative or logistical aspects of the trial.

The Western University NMREB operates in compliance with the Tri-Council Policy Statement Ethical Conduct for Research Involving Humans (TCPS2), the Ontario Personal Health Information Protection Act (PHIPA, 2004), and the applicable laws and regulations of Ontario. Members of the NMREB who are named as Investigators in research studies do not participate in discussions related to, nor vote on such studies when they are presented to the REB. The NMREB is registered with the U.S. Department of Health & Human Services under the IRB registration number IRB 00000941.

Please do not hesitate to contact us if you have any questions.

Sincerely,

Kelly Patterson , Research Ethics Officer on behalf of Dr. Randal Graham, NMREB Chair

Note: This correspondence includes an electronic signature (validation and approval via an online system that is compliant with all regulations).

Appendix Q

Letter of Information

Project Title: Learning from Different Reading Conditions

Principal Investigator

Dr. Frank Boers, PhD

Student Investigator

Liting Luo, MA Student

Faculty of Education

The University of Western Ontario, London, Canada

Thank you for being interested in this research project. Please read the following Letter of Information and decide whether you would like to participate in this project or not. If you decide to participate in this study, we will sincerely appreciate your help. If you decide not to take part in the study, we will also be thankful for your interest.

Invitation to the study project

You are invited to participate in the current study about language learning as you are a potential English language teacher as well as a second language learner. It is expected that you will be in this study for two weeks (two hours in total). It will take approximately 1.5 hour for the first week and half an hour for the second week. The aim of this study is to investigate how reading with different kinds of glosses support language learning.

The rationale of the study

Glosses can be expected to be helpful for second (L2) language learners especially when inferring the meaning of the expression is challenging. A great number of studies have shown a positive effect of glossing on reading comprehension. However, little is known about how to use glossing strategically to maximize its effectiveness for lexical acquisition. Therefore, this research project intends to examine the benefits of reading with different kinds of glosses for language learning. The ultimate goal is to be able to inform language teachers and textbook designers about the kinds of glosses are particularly useful for reading.

The assignment of groups

If you decide to participate then you will be "randomized" into one of

four groups. Randomization means that you are put into a group by chance (like flipping a coin). There is no way to predict which group you will be assigned to. You will have1 in 4 chance of being placed in any group. Neither you nor the researchers can choose what group you will be in. All four groups will use the same reading text, but the glosses in the reading text will be slightly different (but will take the same time).

The procedures of the study

The study consists of 7 stages in total. You are now at stage 0 where you are informed about the study. Those of you who give consent to the participation in the study will proceed to the next stages where you will be asked to read two short English articles (approximately 300 words for each) and perform some tests related to English. This process happens throughout stages 1 to 5. At stage 6, some of you will be invited to have a one-on-one interview. If you do not want to be recorded or interviewed, you can still participate in the other parts of the research. At the final stage of 7, you will be given a debriefing form that explains the purpose of the study in more detail.

The risks and harms of participating in the study

We do not anticipate any risks or discomfort related to participating in this study project, but you may feel tired while completing the activities. However, the researcher will create a comfortable environment, give support, and answer potential questions. The study sessions are well assigned in order to decrease your fatigue.

The benefits of participating in the study project

You are invited to participate in this study because you are a potential English language teacher as well as an English language learner. This study will be beneficial for you as it will allow you to (1) gauge your current vocabulary level, (2) learn several language items, and (3) acquire learning strategies for learning new vocabulary. At the same time, you will be helping with research that is useful for teachers and their students. More specifically, the results of the study will inform learners' and teachers' decision making regarding their use of glosses in reading text.

The option of leaving the study

As your participation in this study is voluntary, you can leave the project at any time. We can also remove your information from the study if you would like us to. If so, you can simply send us an email to let us know of your decision. However, a month after the end of data collection (i.e., after the last session), your data cannot be removed any longer, because we will have started processing the data by then.

Data privacy

All the data collected from you will be kept confidential. We will keep the data for nine years. Only the student investigator and her supervisor (the principal investigator) will have access to the data collected from you, and the data will only be used for the research purposes outlined above. The results of the research project will be reported in a dissertation and possibly in journal articles and conference presentations. No names of any individual students will be mentioned in these reports.

The rights of participants

Your participation in this study is voluntary. You may decide not to be in this study. Even if you consent to participate you have the right to not answer individual questions or to withdraw from the study at any time. If you choose not to participate or to leave the study at any time, it will have no effect on your school grade. You do not waive any legal right by consenting to this study. We will give you any new information that may affect your decision to stay in this study.

Contact for questions

If you have questions about this research study, please contact Dr. Frank Boers, PhD, or Liting Luo, MA Student.

If you have any questions about your rights as a research participant or the conduct of this study, you may contact the Office of Human Ethics. This office oversees the ethical conduct of research studies and is not part of the study team. Everything that you discuss will be kept confidential.

This letter is yours to keep for your future reference.

Appendix R

Consent Form – Student

Project Title: Learning from Different Reading Conditions

Principal Investigator: Dr. Frank Boers

Student Investigator: Liting Luo

For participants

I have read the Letter of Information, and I have understood the nature of the study. All the questions regarding the research project were explained to my satisfaction, therefore, I agree to participate in this research project. I have been provided a copy of the Information Letter and the Consent Form. I know that I may be invited to take part in an audio recorded interview. I voluntarily and freely consent to participate in this study.

_____ _____ _____

Print Name of Participant   Signature   Date (DD-MM-YYYY)

For person obtaining consent

My signature means that I have explained the study to the participant named above. I have answered all questions.

_____ _____ _____

Print Name of Person   Signature   Date (DD-MM-YYYY)

About the results of study:

All the information will be kept confidential to the investigator. If you would like to receive a summary of the research findings in general, please leave your email address here: _____

Obtaining Consent

Appendix S

Debriefing Form

DEBRIEFING FORM

Project Title: Learning from Different Reading Conditions
Principal Investigator: Dr. Frank Boers
Student Investigator: Liting Luo

Thank you for participating in this research project. The purpose of this project was to investigate (1) whether reading with glosses has an effect on L2 idioms learning and (2) which kind of gloss (literal meaning, abstract meaning; abstract plus literal meaning) may help learners better understand and remember L2 idioms. Knowing which kind of gloss is especially useful for learning such expressions will be helpful information for language teachers and textbook designers.

Your results will be kept confidential to the researchers, and all your data will be kept anonymous in any publications. If you have any questions or concerns, please feel free to contact the investigators, Dr. Frank Boers, or Liting Luo.

Here are some references relating to this topic if you want to read more:

Boers, F. (2020). Factors affecting the learning of multiword items. In *The Routledge*

handbook of vocabulary studies (1st ed., pp. 143–157). Routledge.

https://doi.org/10.4324/9780429291586-10

Boers, F. (2021). Glossing and vocabulary learning. *Language Teaching.*

https://doi.org/10.1017/S0261444821000252

Thank you for your participation,
Principal Investigator: Dr. Frank Boers
Student Investigator: Liting Luo
1137 Western Road, London, ON, CA N6G1G7

Curriculum Vitae

Name: Liting Luo

**Post-secondary   Inner Mongolia University of Technology
Education and   Hohhot, Inner Mongolia, China
Degrees:**   2014-2018 B.A.

                 The University of Western Ontario
                 London, Ontario, Canada
                 2018-2019 MPEd.

                 The University of Western Ontario
                 London, Ontario, Canada
                 2020-2022 M.A.

**Honours and   Western Entrance Scholarship
Awards:**   2020-2022

**Related Work   IELTS Teacher
Experience**   Shanghai Global Education Company
                 2020-2022

                 English Language Teacher
                 Likeshuo English for Adults
                 2019-2020

                 English Language Teacher
                 London Language Institute
                 2019