Listening-Based Activities Promoting Productive Knowledge of Formulaic Sequences: Examining the effectiveness of dictogloss

Xi Yu, 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
© Xi Yu 2020

Follow this and additional works at: https://ir.lib.uwo.ca/etd

Part of the Education Commons

Recommended Citation

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact wlsadmin@uwo.ca.
FORMULAIC SEQUENCE DEVELOPMENT

Abstract

Although it is well recognized that second language (L2) learners stand to gain a lot from mastering formulaic sequences (FS) in the target language, research on classroom activities that foster productive knowledge of FSs remains rather scarce. This project, therefore, examines the effectiveness of dictogloss for this purpose. While dictogloss has occurred regularly in research about L2 grammar learning, it has only recently attracted attention from researchers interested in FS learning. This experiment investigated its effectiveness in comparison with two other listening-based activities—dictation and answering comprehension questions. It was hypothesized that dictogloss leads to greater gains in FS knowledge because (a) it involves retrieval of FSs from episodic memory and (b) it makes learners notice (and fill) the gaps in their knowledge.

142 Chinese EFL first-year university students were divided into three treatment groups (dictogloss, dictation, and answering comprehension questions). A pre-test, an immediate post-test, and a delayed post-test were administered to compare the learning gains. Dictogloss was not found to be more effective than dictation for the purpose of FS learning, although both were more effective than answering comprehension questions. Dictogloss did appear to engage students more than dictation with the content of the text. Analyses of the dictogloss worksheets show that it was especially students who successfully retrieved the FSs during text reconstruction who stood a good chance of recalling the FSs in the post-test. This suggests that it is important for teachers who wish to try dictogloss to ensure good accuracy at the reconstruction stage.

Keywords: Dictogloss, formulaic sequence learning, aural modality, retrieval.
FORMULAIC SEQUENCE DEVELOPMENT

Lay People Summary

Learning vocabulary in a second language does not only involve learning a lot of words but also a lot of multiword expressions, referred to as formulaic sequences (FSs). Research has demonstrated that proficiency in a second language and knowledge of FSs are strongly associated. Yet, research on the effectiveness of certain classroom activities intended to foster mastery of multiword expressions remains scarce.

This project examines the potential for FS learning purpose of an activity called dictogloss, in which students listen to a short text a few times, try to reconstruct it from memory, and then compare their version with the original text. Dictogloss has figured rather prominently in research about learning and teaching grammar, but it has only recently attracted attention from researchers interested in learning and teaching FSs. Only two studies have illustrated its potential for FS learning so far, but these did not compare the effectiveness of dictogloss with other activities. This experiment, therefore, investigated its effectiveness in comparison with two common classroom listening activities—dictation and answering listening comprehension questions.

142 Chinese EFL first-year university students were divided into three groups (dictogloss, dictation, and answering comprehension questions). They completed a number of tests to compare how many FSs were learned as a result of these three listening-based activities. The results suggest that dictogloss was not more effective than dictation for the purpose of FS uptake, although both dictogloss and dictation were more effective in this regard than answering comprehension questions about the text. On the positive side, dictogloss appeared to engage students with the content of the text to a degree that is comparable to answering comprehension questions, whereas dictation did not.
FORMULAIC SEQUENCE DEVELOPMENT

Comparisons of the FSs on the dictogloss worksheets and the test sheets reveal that it is especially when students successfully recalled FSs as they reconstructed the text that they stood a good chance of recalling the FSs also in the subsequent test. This finding, therefore, suggests that it is important to implement dictogloss activities in ways that ensure a high degree of success when the students try to reconstruct the text.
Acknowledgements

First of all, I would like to express my sincere appreciation to my supervisor, Dr. Frank Boers. Had it not been for your excellent supervision and mentorship, this thesis would never have come into being. I have learnt immensely through your insightful feedback and timeless support. Thank you for being the most patient listener and supporter.

I would also love to highly appreciate the enthusiastic engagement of students, teachers, and deans at Beijing International Studies University, China. Without their great support, it would have been impossible for me to finish this massive data collection.

I thank my seniors and friends of the vocabulary group at the Faculty of Education of Western University. They have both inspired and supported me along the way.

Heartfelt thanks also go to my parents and grandparents for always being with me and constantly encouraging me.
# Table of Contents

Abstract ......................................................................................................................... i

Lay People Summary ................................................................................................. ii

Acknowledgements ....................................................................................................... iv

Table of Contents .......................................................................................................... v

List of Tables ................................................................................................................. viii

List of Figures ............................................................................................................... ix

List of Appendices ........................................................................................................ x

Chapter 1 Introduction ............................................................................................... 1

Chapter 2 Literature Review ....................................................................................... 3

  2.1 Definition of Formulaic Sequence ........................................................................ 3

  2.2 The Significance of FS Knowledge ....................................................................... 5

  2.3 The Origins of FS Learning Difficulties .............................................................. 7

  2.4 Interventions to Promote FS Learning ............................................................... 9

  2.5 Dictogloss ............................................................................................................. 15

  2.6 Dictogloss to Facilitate Grammar and Vocabulary Learning ............................. 17

  2.7 Dictogloss to Promote FS Learning .................................................................... 19

  2.8 Previous Studies on Dictogloss for FS Learning .................................................. 21

Chapter 3 Rationale for the Research Design .............................................................. 23

Chapter 4 Research Questions ..................................................................................... 25
FORMULAIC SEQUENCE DEVELOPMENT

Chapter 5 Method ........................................................................................................... 27
  5.1 Research Design ..................................................................................................... 27
  5.2 Participants ............................................................................................................. 28
  5.3 The Input Text ....................................................................................................... 30
  5.4 Target Items .......................................................................................................... 31
  5.5 Pre-teaching and warm-up .................................................................................... 33
  5.6 Procedures ............................................................................................................ 34
  5.7 Instruments .......................................................................................................... 40

Chapter 6 Data Processing ............................................................................................ 42

Chapter 7 Analysis ........................................................................................................ 44

Chapter 8 Results ......................................................................................................... 45

Chapter 9 Discussion ................................................................................................... 51

Chapter 10 Conclusions and Pedagogical Implications ............................................... 55

Chapter 11 Limitations and Directions for Further Research ..................................... 57

References ..................................................................................................................... 60

Appendices ..................................................................................................................... 75

  Appendix A: Ethics Approval Notice ......................................................................... 75
  Appendix B: Letter of Information-Student ................................................................. 76
  Appendix C: Consent Form .......................................................................................... 80
  Appendix D: A sample of the Updated Vocabulary Levels Test ................................ 81
FORMULAIC SEQUENCE DEVELOPMENT

Appendix E: Pre-test .................................................................................................................. 82
Appendix F: Listening Material ................................................................................................. 84
Appendix G: Warm-up discussion .............................................................................................. 85
Appendix H: Immediate Post-test ............................................................................................. 86
Appendix I: Delayed post-test .................................................................................................. 88
Appendix J: Questions for Listening Comprehension Group .................................................... 90
Curriculum Vitae ....................................................................................................................... 91
# List of Tables

Table 1 ................................................................................................................................. 29
Table 2 ................................................................................................................................. 30
Table 3 ................................................................................................................................. 33
Table 4 ................................................................................................................................. 34
Table 5 ................................................................................................................................. 35
Table 6 ................................................................................................................................. 36
Table 7 .................................................................................................................................. 46
Table 8 .................................................................................................................................. 46
Table 9 .................................................................................................................................. 46
List of Figures

Figure 1. Different components of the research design ................................................................. 27

Figure 2. The Estimated Mean of Providing a Correct Response on the Immediate Posttest and Delayed Posttest as a Function of Treatment .................................................................................. 48
List of Appendices

Appendix A: Ethics Approval Notice ................................................................. 75
Appendix B: Letter of Information-Student .................................................... 76
Appendix C: Consent Form .............................................................................. 80
Appendix D: A sample of the Updated Vocabulary Levels Test ...................... 81
Appendix E: Pre-test ......................................................................................... 82
Appendix F: Listening Material ....................................................................... 84
Appendix G: Warm-up discussion .................................................................. 85
Appendix H: Immediate Post-test .................................................................... 86
Appendix I: Delayed post-test ......................................................................... 88
Appendix J: Questions for Listening Comprehension Group ........................... 90
Chapter 1 Introduction

For language learners, the ultimate goal is using the language fluently and accurately. Research has indicated that knowledge of conventional multiword expressions, often referred to as formulaic sequences (FSs) in the literature, positively influences learners’ language proficiency (e.g., Boers, Eyckmans, Kappel, Stengers, & Dememcheleer, 2006; Crossley, Salsbury, & McNamara, 2015; Kremmel, Brunfaut, & Alderson, 2017; Saito, 2020; Tavakoli & Uchihara, 2019). However, it is difficult for L2 learners to acquire FS knowledge from mere exposure to natural discourse, especially for learners in an English as a foreign language (EFL) context. Hence, in the last two decades, researchers have attempted to seek effective and efficient treatments to foster this knowledge. With respect to formulaic sequence recognition, many interventions have been shown to be effective (e.g. Peters & Pauwels, 2015). Unfortunately, promoting productive knowledge of formulaic sequences appears to be much harder.

This thesis examines the potential benefits for FS learning of a classroom activity called dictogloss (Wajnryb, 1990). This is an activity where students listen to a text several times, note down keywords, and then try to reconstruct the text from memory. Although it was originally proposed as an activity to raise students’ awareness of grammar features (in fact, it has also been referred to as “grammar dictation”; Wajnryb, 1990), it has recently attracted the interest of researchers interested in its potential for FS learning (e.g., Lindstromberg, Eyckmans & Connabeer, 2016). As will be discussed in more detail below, one of the reasons for believing dictogloss could be beneficial for FS learning is that it involves retrieval from memory (since learners are asked to reproduce a text they have listened to).

Text-reconstruction activities such as dictogloss have been researched mostly with a focus on grammar learning (e.g., Izumi, 2002; Li, Ellis & Zhu, 2016), while only very few
studies have explored their merits for FS learning. The classroom experiment reported in this thesis therefore examined the potential of the dictogloss activity in fostering productive knowledge of formulaic sequences by comparing it with two other, more common, listening-based activities—a regular dictation and answering listening comprehension questions. The findings of the study may be useful for teachers and materials designers looking to expand their arsenal of FS-oriented exercises and activities.

The experiment, which was conducted with the participation of EFL students in mainland China, received approval of the Western University Non-Medical Research Ethics Board. The approval document is provided in Appendix A.
Chapter 2 Literature Review

This part begins with defining the type of multiword items this project focuses on. Then, the essential role of formulaic sequence knowledge in second language learning is pointed out by discussing the ubiquity of FSs in natural discourse and the contribution of FS knowledge to learners’ language proficiency. In the third part, four challenges that make FS acquisition problematic for L2 learners are illustrated. Pedagogical interventions to promote FS learning that have been tried so far are reviewed in the fourth part. Finally, the potential benefits of the dictogloss activity, which has so far attracted little attention from researchers seeking ways of fostering FS knowledge, will be introduced.

2.1 Definition of Formulaic Sequence

The last 20 years have seen an increased interest in the methods used to learn and teach multiword items (for reviews, see Boers and Lindstromberg, 2012, Boers and Webb, 2018, and Pellicer-Sánchez and Boers, 2018). Researchers have defined multiword items in their own ways, leading to some disparities in the definition as it appears in the literature, and so it is necessary to define “multiword items” in this study.

Wray (2002) proposed the term “formulaic sequence”, which she defined as:

a sequence, continuous or discontinuous, of words or other elements, which is, or appears to be, prefabricated: that is, stored and retrieved whole from memory at the time of use, rather than being subject to generation or analysis by the language grammar. (p. 9)

In this definition, Wray (2002) emphasises two features. First, the definition points out the above-single-word nature of FSs. Second, the definition highlights the prefabricated nature of
formulaic sequences, stating that they appear to be stored in and retrieved from memory as whole units. Additionally, “formulaic” of course implies that these expressions are highly conventional in the given language community.

**Using “formulaic sequence” (FS) as the umbrella term in this study**

There are many different types of multiword items, including (1) collocations (e.g. gain weight), (2) idioms (e.g. jump the gun), (3) phrasal verbs (e.g. give away), (4) compounds (e.g. press conference), (5) binomials (e.g. more or less; fish and chips), (6) lexical bundles (e.g. as a result of), (7) standardized similes (e.g. good as gold), (8) proverbs (e.g. The early bird catches the worm), (9) conversational formulae (e.g. How are you?), (10) aphorisms (e.g. actions speak louder than words), (11) lexicalized sentence stems (e.g. chances are), and other readymade units.

Authentic texts usually include various kinds of multiword items, but it would not be very practical to use a specific term to characterize each instance. Hence, an umbrella term is useful. In this dissertation, I will adopt formulaic sequence (FS) as the umbrella term, while fully recognizing this represents diverse categories of multiword items, each of which may pose specific challenges for learning (Boers, 2020). I also recognize that Wray’s (2002) suggestion that FSs are processed “holistically” applies primarily to L1 users. It does not necessarily hold true for post-childhood L2 learners, who are probably more likely to process such sequences word by word in the early stages of becoming familiar with them. Still, as they become increasingly familiar with a certain FS, also L2 learners will eventually be able to process and produce it effortlessly and without hesitation, and so this will at least give the phrase the appearance of being prefabricated.
2.2 The Significance of FS Knowledge

Cowie (1992) points out that “it is impossible to perform at a level acceptable to native users, in writing or in speech, without controlling an appropriate range of multiword units” (p.10). This section aims to demonstrate the significance of formulaic sequence knowledge for language learners by pointing out the ubiquity of FSs in natural language and their essential role in learners’ language proficiency, along three dimensions: (a) fluency, (b) complexity and (c) accuracy (Boers et al., 2006). The following section will discuss these dimensions separately.

Ubiquity. A substantial proportion of natural spoken and written discourse is made up of FSs (Conklin & Schmitt, 2012; Erman & Warren, 2000; Foster, 2001; Hill, 2000; Pawley & Syder, 1983; Shin & Nation, 2007). Based on different definitions of FS, the estimated proportion of FSs in natural discourse ranges from 32% (Foster, 2001) to more than 50% (in spoken discourse) (Erman & Warren, 2000). This omnipresence of FSs in natural discourse makes them essential for language learners to learn.

The contribution of FSs to learners’ language fluency. A considerable number of publications indicate that FS mastery is a key facilitator for language learners’ receptive as well as productive language fluency (e.g., Pawley & Snyder, 1983; Wray, 2002). For receptive fluency, FS knowledge makes discourse more predictable. Anticipating formulaic language allows listeners or readers to pay more attention to non-formulaic items, thus aiding processing and comprehension (Boers & Lindstromberg, 2012). For example, on hearing or reading on the one hand, high proficiency learners are likely to expect on the other hand in the following part of the text and anticipate “hand” after hearing or reading “on the other”.

When it comes to productive fluency, Boers et al. (2006) and Stengers, Boers, Housen, and Eyckmans (2010, 2011) found that learners who use more FSs in their language were rated
as relatively fluent speakers. A similar result was found in studies by Wood (2006, 2010), Ushigusa (2008), and more recently Tavakoli and Uchihara (2019). Learners with good mastery of a large FS repertoire can speed up their speaking and writing through retrieving multiword items from their memory as prefabricated units, instead of assembling them word by word at the time of speaking or writing.

Learners whose discourse use exhibits good mastery of FSs can use these expressions without much effort, and this frees up attentional resources which can be devoted to the content of the text (Coxhead, 2007), which thus not only helps learners express themselves well but also helps them comprehend texts better (e.g., Kremmel et al., 2017).

**The influence of FS knowledge on learners’ language complexity.** In the applied linguistics field, complexity nowadays is understood to include not only syntactic complexity but also lexical richness or lexical sophistication. L2 discourse rich in FSs tends to be viewed by assessors as exhibiting a wide range of expression (Boers et al., 2006; Saito, 2020).

**The role of FS knowledge for learners’ accuracy.** A lack of FS knowledge leads learners to erroneously devise non-standard or non-native like expressions (Cobb, 2018). Good knowledge of FSs promotes learners’ language accuracy also “because these prefabricated chunks constitute ‘zones of safety’ and appropriate use of them may thus confine the risk of ‘erring’ to the spaces in between the formulaic sequences in one’s discourse” (Boers et al., 2006, p. 247).

To sum up, research indicates that FS knowledge is related to learners’ language proficiency (Bestgen, 2017; Bartning, Forsberg & Hancock, 2009; Crossley, Salsbury, & McNamara, 2015; Dai & Ding, 2010; Forsberg, 2010; Granger & Bestgen, 2014; Kremmel,
Due to the ubiquity and its contribution to learners’ proficiency, FS knowledge is indispensable for language learners. Unfortunately, it appears to be hard for learners to acquire FSs from exposure to natural discourse alone. The following section discusses some of the challenges of FS acquisition.

2.3 The Origins of FS Learning Difficulties

Researchers agree that it is hard for post-childhood language learners to acquire and use FSs, especially in contexts where exposure to L2 discourse is limited, such as in English as a Foreign Language (EFL) contexts (e.g., Granger, 1998; Laufer & Waldman, 2011; Li & Schmitt, 2009; Nesselhauf, 2003; Siyanova & Schmitt, 2008; Siyanova-Chanturia & Spina, 2015). Research has found that even advanced language learners use a narrower range of FSs than native speakers/writers do, and that they still misuse FSs in their target language (Altenberg & Granger, 2001; Laufer & Waldman, 2011; Nesselhauf, 2005). The following section discusses five possible origins of learners’ FS acquisition difficulties.

First, FSs may consist of highly frequent words (e.g., *make mistakes and *in other words), which, owing to their familiar appearance, tend not to attract much attention (Godfroid, Boers, & Housen, 2013; Tulving & Kroll, 1995). As a result, learners may overlook their phraseological or collocational patterning, which, in turn, may lead to erroneous substitutions or omissions of such words (e.g., *do mistakes and *with other words).

Second, identifying FSs in natural discourse is far from easy. While single words are demarcated in print thanks to the empty spaces between them, this does not help to demarcate
FORMULAIC SEQUENCE DEVELOPMENT

FSs (Lindstromberg, Eyckmans & Connabeer, 2016). Besides, morphological and syntactic variability can also hinder learners’ recognition that certain words form strong lexical partnerships (Boers, 2020).

Third, although a high proportion of natural discourse consists of FSs, few individual FSs occur repeatedly in the same text. For example, Boers and Lindstromberg (2009) found only one example of *tell the truth* in 100 pages of a police detective story, while this type of text could be expected to contain a relatively high number of occurrences of this expression. Beyond a small number of highly frequent FSs (including fillers such as “kind of” and “you know”), learners who are not exposed to an abundance of L2 input can thus not be expected to encounter the same FS several times in a short time span. This is relevant because research has shown that learners usually need multiple encounters with the same FS for measurable learning to happen (see below). Learners in an EFL context have limited opportunities to encounter FSs outside their textbooks and classroom interactions (including teacher talk), and so they are unlikely to encounter a wide range of FSs, let alone encounter one and the same FS multiple times.

Fourth, many FSs express a meaning that does not follow straightforwardly from the constituent words (Martinez & Murphy, 2011). Owing to their puzzling semantics, such expressions (commonly referred to as idioms in the literature on phraseology) may perhaps attract some more attention from learners than totally transparent phrases, but, on the downside, it will take extra steps to work out their meaning even when supportive context is available (Boers, Eyckmans, & Stengers, 2007).

Fifth, research has found that language teachers and learners tend to prioritize single word learning over FS learning (Bui, Boers & Coxhead 2019; Nguyen, 2014; Peters, 2012).
Their underestimation of the importance of learning FSs may be an additional explanation for the slow pace of FS acquisition by L2 learners.

### 2.4 Interventions to Promote FS Learning

As acquisition of FSs from mere exposure to unmodified natural discourse is likely to be a very slow process (especially in EFL contexts), it is worth investigating the effectiveness of pedagogical interventions. In literature, such interventions are usually classified into two broad categories: incidental learning interventions and deliberate learning interventions. Incidental learning refers to language learning as a by-product of communication (Ellis, 1999; Schmitt, 2010), whereas deliberate learning involves a learner’s intention to study and remember linguistic elements or features. On the one hand, it has been argued that deliberate learning generates noticeable learning gains the fastest (e.g. Laufer, 2003; Lin & Hirsh, 2012). On the other hand, incidental learning conditions by their very nature leave more space for developing language skills (e.g., reading and listening skills) alongside the acquisition of discrete language items or features (Nation & Webb, 2011).

In practice, however, incidental learning and deliberate learning are not dichotomous because it is hard to identify whether (and when) during communication a learner’s attention is on the language code or on the content of the messages. Nation and Webb (2011) propose to operationalize deliberate learning conditions as conditions where learners are explicitly told to focus on language items. While this provides a useful way of distinguishing deliberate from incidental learning, some scholars have argued that finer-grained distinctions might be useful within the realm of incidental learning. For example, if learners read a text with textually enhanced language elements, the purpose of the enhancement is to attract the learners’ attention,
and so learners may become aware that they are meant to read the text not only for its content but also to take notice of the enhanced language elements. In their review of pedagogical approaches to formulaic language, Pellicer-Sánchez and Boers (2018) propose an additional category besides ‘purely’ incidental and deliberate learning conditions, which (for lack of a better term) they label “semi-incidental”. These are conditions which direct learners’ attention to particular FSs in a text, while the learners are nonetheless expected to engage with the text first and foremost for its content.

The next three sections examine the three types of conditions for FS learning in turn.

**Incidental Learning Conditions.** When it comes to incidental learning treatments, several studies have investigated the effects of (a) frequency of encounters and (b) modality of input. These two factors have been found to matter for single word learning, and so there are grounds for assuming they also matter for FS learning.

An early investigation into repetition was conducted by Durrant and Schmitt (2010). In their study, participants were assigned to one of three conditions: (1) target collocations repeated once in a sentence context; (2) repetitions of the same sentences containing the target collocations; and (3) the target collocations embedded in different sentences. The participants in both repetition conditions were found to have better retention of the target collocations than their peers in the non-repeated condition. Webb, Newton and Chang (2013) embedded Verb + Noun collocations in texts, with different numbers of encounters (1, 5, 10, 15 times). The best learning gain was observed for the 15-encounters group. However, it is worth mentioning that even as many as 10 encounters did not ensure correct recall of several of the target collocations. Five encounters did not generate much greater gains than a single encounter at the level of productive recall (also see Pellicer-Sánchez, 2017). Still, “seeding” a text with numerous instances of the
same FSs clearly improves the chances of learners picking them up. It needs to be mentioned, however, that incorporating multiple instances of the same FSs in a text requires a fair amount of effort and creativity on the part of materials writers, and many busy teachers may shy away from trying this themselves.

Modality of input is another variable. For example, reading a text while listening to its audio-recording appears to lead to better vocabulary uptake than reading or listening alone (e.g., Teng, 2018), and this holds true also for FSs (Webb & Chang, under review). The use of audio-visual input (video) has attracted a lot of attention in recent years (e.g., Peters & Webb, 2018), and a recurring finding in this strand of research is that the provision of captions facilitates vocabulary uptake, too (Montero Perez, Van Den Noortgate, & Desmet, 2013; Montero Perez, Peters, Clarebout, & Desmet, 2014). Again, there is some evidence that this positive effect of captions extends to FSs (Majuddin, Siyanova-Chanturia, & Boers, 2019; Puimège & Peters, 2019).

At the same time, additional steps may be required if the aim is for learners to add new FSs to their own L2 repertoire because learners seem to have little inclination to spontaneously recycle language from an input text beyond single content words. This, at least, is what was observed in a study by Hoang and Boers (2016) in which ESL students twice read and listened to an illustrated story. The participants were instructed that they would be asked to retell the story immediately afterwards. Although the story contained many FSs, less than 7% of these FSs were accurately used in the participants’ story retelling. The participants did recycle many of the content words from the input text, but without their phraseological patterns (e.g., request but not make a request; notice but not take no notice of). It is worth mentioning here that the participants had not been explicitly encouraged to recycle the language used in the original story in their own
retelling of the story. Neither were they asked afterwards to compare their own rendering of the story with the original text. As we shall see further below, this is different from the intervention to be examined in the present research project—the dictogloss activity.

**Semi-incidental Learning Conditions.** According to the Noticing Hypothesis (Schmidt, 2001), “people learn about the things they attend to and do not learn much about things they do not attend to” (p. 30). We should keep in mind that learners do not always pay attention to unknown linguistic features and do not always realize they do not yet know them (Laufer, 2005). Therefore, researchers have explored ways to direct learners’ attention to linguistic features in texts but without compromising their intake of the content of texts. Learning resulting from this is labelled “semi-incidental” in Pellicer-Sánchez and Boers’ (2018) review.

Under semi-incidental learning conditions, most of the investigations focus on typographic enhancement, an intervention originally proposed to improve grammar learning (e.g., Indrarathne & Kormos, 2017; Jahan and Kormos, 2015; Lee, 2007; Winke, 2013). Just a few studies to date have explored the benefits of typographic enhancement as a means of promoting FS learning. In a study reported by Boers, Demcheleer, He, Deconinck, Stengers, and Eyckmans (2017), participants read texts with underlined FSs. Post-test results showed that learners were more likely to recognize the FSs they had seen with underlining than ones they had encountered in the text without underlining. Choi (2017) also demonstrated that typographic enhancement (boldface) attracts attention (as gauged through eye-tracking) and leads to better recall of enhanced collocations. Szudarski and Carter (2016) compared input flooding alone (i.e., incorporating multiple instances of the same FSs) with input flooding combined with typographic enhancement, and found that the second condition led to better FS learning. On the downside, research has also found that there is a risk of a “trade-off” in learners’ intake of
enhanced and unenhanced elements from a text (Boers et al., 2017; Choi, 2017), suggesting that attention-directing through typographic enhancement may come at some cost.

Typographic enhancement can also be applied to captions in videos to help FSs stand out, but the effect of this seems relatively small (Majuddin et al., 2019), possibly because viewing a video requires processing that is too fast for learners to dwell long enough on the typographically enhanced word sequences in the captions.

**Intentional Learning Conditions.** It is widely believed that intentional vocabulary learning brings about more noticeable gains in a short period of time than incidental learning conditions (e.g., Laufer, 2003), and this belief includes FS learning (Laufer & Girsai, 2008; Nesselhauf, 2003, p. 238; Szudarski & Carter, 2014). However, while some explicit pedagogical procedures and exercises tend to be effective with respect to FS recognition and the development of strategies for FS learning, their effects on productive knowledge of FS are not always as clear.

In Boers et al. (2006), participants in the experimental group were asked to pay attention to the target FSs in the listening and reading texts used in their course and to regularly do the “text-chunking activity” (Lewis, 1997), where they looked for FSs in the texts under the guidance of the teacher. The participants in the comparison group were not alerted to the FSs in the texts. After a school year, an interview was conducted to gauge learners’ oral proficiency. The result showed that the learners in the experimental group used more FSs in their interview than the learners in the comparison group. However, the FSs used in the interview were mostly expressions from a new text provided to the students to talk about in the interview. Therefore, although learners seemed equipped with FS learning strategies after the intervention (they spontaneously recycled FSs from the new text), there was no compelling evidence that they had in fact learned many FSs from the texts used in their course. In a conceptual replication, Bui et
al. (2019) administered tests specifically on the FSs included in the course texts and they did find evidence that the experimental group recalled these better than the comparison group. However, this effect of the intervention was on the small side. It is worth mentioning here that the class activities in these two studies where the students were invited to engage with the phraseological units in the texts were not followed by output activities which invited the students to use these expressions.

Indeed, receptive activities seem not to lead learners to incorporate many newly encountered FSs in their own L2 output (Peters & Pauwels, 2015). One explanation is that learners find it hard to focus on the accurate use of newly learned items as their attentional resources are directed mostly to the content of their discourse rather than its wording. Another explanation is that productive retrieval practice is desirable to transfer receptive knowledge into productive knowledge (e.g., Webb, 2005; Zhang, 2017). If so, FS-focused exercises might generate clearer evidence of learning gains. For example, Laufer and Girsai (2008) found their EFL students remembered target collocations better after doing focused exercises on them (especially translation exercises). However, the format of such FS-focused exercises matters, and there are indications that certain formats (such as exercises that require students to re-assemble broken-up collocations) bring about only marginal learning gains (Boers, Demecheleer, Coxhead, & Webb, 2014; Boers, Dang, & Strong, 2017). It has also been cautioned that such exercises should be implemented as retrieval practice, not as trial-and-error events (e.g., Strong & Boers, 2019a/b). In other words, learners should first be exposed to the correct FSs and then recall these from memory instead of being asked to guess the composition of FSs, because wrong guesses risk causing interference later on, especially in the case of constituent words with poor
semantic and/or formal distinctiveness (e.g., do vs. make; make vs. take; in vs. on) and in the case of near-synonyms (e.g., say vs. tell; hold vs. keep; tall vs. high) (Webb & Kagimoto, 2011).

To sum up, research on incidental and semi-incidental learning has been prevalent in this field because it provides opportunities for learners to develop skills beyond discrete language item learning. Incidental learning is a rather slow process, however. Deliberate language learning activities may generate learning gains faster, but their observed benefits have not been particularly compelling either. In the case of text-based noticing activities (e.g., Boers et al. (2006), evidence of an effect on productive recall has not been very compelling either. In the case of decontextualized exercises, the evidence has been mixed. So, especially when it comes to promoting productive FS knowledge, additional research on pedagogic interventions would clearly be welcome. The following section discusses such an intervention, called “dictogloss”, which (for reasons discussed below) appears in theory to have good potential for fostering productive FS knowledge.

2.5 Dictogloss

Dictogloss was introduced by Wajnryb (1990) as an adapted dictation task to promote grammar learning. The dictogloss activity consists of four steps (Wajnryb, 1990: p.7):

a. **Preparation:** The teacher introduces the topic of the text and administrates a warm-up activity to familiarize the students with the topic, so as to facilitate their comprehension of the text they will listen to.

b. **Dictation:** The teacher reads the text twice at a natural pace. The first time, students are instructed to listen without taking notes; the second time, they are encouraged to write down key words that will help them recall the content of the text.
c. **Reconstruction:** Students work in small groups to share their notes and to reconstruct the original text.

d. **Analysis and Correction:** There are various ways to analyse the students’ written products and to give feedback, and teachers could choose the one that they find most suitable. The students will typically be given the original text and asked to detect how their own version is different from the “model”. This can be done individually, collaboratively or even as a whole-class activity.

I implemented the dictogloss activity slightly differently in my classroom experiment. First, instead of reading the English text aloud myself, I used an audio-recording by an L1 user. Not only did this ensure appropriate pronunciation and prosody, but it also ensured that the pace of speech would be identical for each of the treatment groups in the experiment. Second, piloting revealed that listening to the recording only twice was insufficient for students to write down enough keywords to help them successfully recall the text content, and so I decided to play the recording four times. This also allowed the students to write down more of the content words that happened to be constituents of FSs, which put the students in a better position to reproduce the complete FSs in the subsequent text-reconstruction stage. Third, before reconstructing the text, the students were given several minutes to re-organize their notes and try to insert some missing words individually. Then, the students worked in pairs and shared their notes to reconstruct the text. Pair work was considered more practical than groupwork, as it did not require changing the conventional seating arrangement in the (Chinese) classroom. It is also more practical to compare two worksheets than three or four worksheets at the same time. After having produced their own version of the text, the students were asked to compare this to the
original text, to discuss the differences, and make amendments to their own version using a pen or a pencil with a different color.

Dictogloss may be considered an incidental learning condition to the extent that learners need to engage first and foremost with the content of the text to be able to reconstruct it. However, the students are also instructed that their final product should resemble the original text as closely as possible, and so this would at some stage of the activity require a focus on language proper. In general, a dictogloss activity is relatively easy for teachers to implement in their classrooms, and it provides an additional way of using listening input besides more familiar and common activities such as listening with a view to answering comprehension questions.

2.6 Dictogloss to Facilitate Grammar and Vocabulary Learning

Text reconstruction activities such as dictogloss have been used in many studies to prompt so-called language-related episodes in learner-learner interaction. These are episodes where the learners discuss how something is said in the target language rather than what is said. Observations of language-related episodes triggered by dictogloss (or similar activities) played a major part in the development of Swain’s Output Hypothesis (Kowal & Swain, 1994, 1997; Swain, 1998; Swain & Lapkin, 1998; 2002), and were later integrated under the broader notion of “Languaging” (e.g., Swain, 2010). When students collaborate to reconstruct a text from memory, they become aware of lacunae in their linguistic resources and may fill these lacunae through seeking help from each other and through revisiting the original text (e.g., Bastarrechea, García Mayo, & Leeser, 2014; Gallego, 2010; Kuiken and Vedder, 2002; Leeser, 2004; Uludag & Vanpatten, 2012; Vanpatten, Inclezan, Salazar, & Farley, 2009; Williams & Evans, 1998). Some studies have demonstrated significant learning gains from this in the realm of grammar.
acquisition (e.g., Izumi, 2002; Li, Ellis & Zhu, 2016; Russell, 2014; Song and Suh, 2008).

Dictogloss has been recommended in various practitioner-oriented resources as well (e.g., Smith, 2012; Thornbury, 2006). There are, however, also a few empirical studies where dictogloss was not found to prompt much attention to grammar features (García Mayo, 2002; Gorman & Ellis, 2019; Nabei, 1996). García Mayo (2002), for instance, found far fewer language-related episodes during a dictogloss activity than during an activity where students were asked to find and correct mistakes in a text. This study did not examine actual learning outcomes, however, and so should not necessarily be taken as evidence in favor of the latter type of activity. Some researchers have called for caution with the use of “find and correct errors” exercises because they present students with mistakes that they themselves might never have made and which may interfere later with recall of the correct alternatives (Boers & Strong, 2016). Gorman and Ellis (2019) implemented a series of dictogloss activities to examine young learners’ development in their accurate use of English tenses and found little improvement over time.

Dictogloss arguably has greater potential for lexical development, because language-related episodes arising in dictogloss activities often concern lexis (e.g. Dobao, 2014; García Mayo, 2002; Nabei, 1996). It is surprising, then, that few studies have explored the use of dictogloss to foster vocabulary learning. One of the exceptions is a study by Kim (2008), where 32 students of Korean did a dictogloss activity either individually or collaboratively, and which found better performance on a vocabulary post-test after the collaborative condition. This was a comparison of different implementations of dictogloss, however, and so does not as such demonstrate that dictogloss is more effective for vocabulary learning than other text-based activities. To my knowledge, there is hardly any research which compares dictogloss with other listening-based activities (such as regular dictation). One exception is a small-scale study by
Reinders (2009), which showed a small advantage of dictogloss over a regular dictation for students’ acquisition of subject-verb inversion in phrases such as *Seldom had he + past participle* and *No sooner had he + past participle*.

Of interest as well is the question whether L2 students themselves experience dictogloss as a useful activity. This question is especially pertinent because implementing a dictogloss activity can be quite time consuming. Research by Gallego (2014) is reassuring in this regard, because the (approximately 500) learner-respondents in her study were, overall, in the opinion that dictogloss was useful and effective for language learning, although it was especially the comparatively advanced learners whose reactions were positive.

### 2.7 Dictogloss to Promote FS Learning

As there is evidence that dictogloss is beneficial for grammar learning and most probably also for word learning, it seems worth expanding its scope of application to FS learning. There are several reasons for believing that dictogloss can be beneficial for FS learning.

First, dictogloss is a listening-based activity. It was mentioned above that identifying FSs in written texts is difficult for language learners because there are few visual cues to demarcate where a multiword expression starts and where it ends. A listening activity, on the contrary, involves prosody, which supports the parsing of text and may help learners to identify semantic units such as FSs (Lin, 2012). For example, speakers are very unlikely to pause within a formulaic sequence. Rather, pauses will often signal the boundaries of multiword lexical units.

Second, since the students listen to the same text several times in a dictogloss lesson, they will also be exposed to the same FSs several times. As mentioned above, learners are unlikely to encounter the same FS in a single text several times. Research has shown the benefits of story
repetition for word learning (Elley, 1989; Elley & Mangubhai, 1983). Listening to the same text several times creates repetition that would otherwise be lacking. It may be argued that there is a downside to this, because the same FSs are encountered in the same context. Perhaps varied contexts would be more helpful for learners to develop a good understanding of the meaning and usage patterns of the expressions. On the other hand, according to a study by Durrant and Schmitt (2010), which I reviewed above, re-encounters with FSs in the same context might in fact be beneficial in the early stages of learning a new FS. Re-engagement with the same text frees up cognitive resources—the first time one will tend to devote most resources to comprehending the content (*what* is said), but afterwards resources can be allocated to the linguistic packaging of the content (*how* it is said).

Third, dictogloss pushes output, and it is widely accepted that output is a key factor of language learning—especially for the development of productive knowledge. The benefits of output are well explained by Swain and Lapkin (1995):

In producing the L2, a learner will on occasion become aware of (i.e., notice) a linguistic problem (brought to her his/ her attention either by external feedback (e.g., clarification requests) or internal feedback. Noticing a problem ‘pushes’ the learner to modify his/her output. (pp. 372–73)

In the case of dictogloss, students are likely to compare their own output attempts to how the same message is worded in the model text, especially where they felt unsure of the accuracy of their own wordings. This, then, will present an opportunity for adjusting their own L2 resources. In a similar vein, the discussion between students as they work in pairs to reconstruct the text may help them to adjust their interlanguage even before seeking confirmation from the original text. The student-student task-related questions and comments about language may well
include questions and comments about the composition of FSs, including the accurate choice of function words (or closed-class words) such as prepositions (Nabei, 1996).

Fourth, because students will unlikely be able to write down full FSs as they listen to the text during the note-taking stage, they will need to try and complete them from memory. They may have written down certain semantically distinct and relevant content words, but not the less distinct accompanying elements (or collocates) such as delexicalized verbs, prepositions and articles. These will thus need to be added from memory. It is widely accepted that retrieval is important for making new knowledge stick (Karpicke, Lehman, & Aue, 2014). For example, showing learners new word-meaning pairings and then asking them to recall these leads to stronger memories than merely asking learners to study and re-study the pairings (e.g., Barcroft, 2007).

Finally, dictogloss requires learners to write down words and phrases, and it has been suggested that word writing facilitates vocabulary learning. Candry, Elgort, Deconinck, and Eyckmans (2017), for example, found that word-writing yielded higher post-test scores than a meaning-inferencing condition. Webb and Piasecki (2018) also found that writing down words led to better retention than a paired-associate learning condition which did not involve writing.

2.8 Previous Studies on Dictogloss for FS Learning

Although dictogloss seems to have good potential to promote knowledge of FSs, to my knowledge, just three published studies (Lindstromberg et al., 2016; Snoder & Reynolds, 2018; Wood, 2009) have empirically investigated its effectiveness for this purpose.

Dictogloss was one of several activities in a fluency workshop with a focus on FSs by Wood (2009), who describes one ESL learner’s improvement in speech fluency and its
association with FS use. Since dictogloss was just one of many activities implemented in this case study, it is impossible to tell precisely how much it contributed to the learner’s progress. Lindstromberg et al. (2016) compared a modified version of dictogloss, where the targeted FSs were given on the work sheets, with the standard dictogloss (a blank sheet of paper was given to learners to reconstruct the text). The modified dictogloss group showed better learning than the standard dictogloss group. This is somewhat surprising because the modified dictogloss did not require learners to retrieve the FSs from memory (as these were provided on the worksheet) and, as mentioned, retrieval is generally known to be beneficial. One possible explanation is that presenting the target FSs on the worksheet ensured these were included accurately in the students’ final version of the text (and thus without a risk of initial erroneous representations). It is also possible that their presence on the worksheet alerted the students to the precise language focus of the classroom experiment.

Snoder and Reynolds’s (2018) study was inspired by Lindstromberg et al. (2016). This within-group study investigated the extent to which different pre-task activities followed by a modified dictogloss activity affected verb-noun collocations learning. Participants first received a glossed list of target collocations 1-6, generated original sentences with the target collocations and then did the first dictogloss activity. The same group of participants then received target collocations 7-12 and were asked to do a rhyming activity. The authors found better learning gains for the collocations which the students had invented new sentences for, presumably because this activity requires semantic processing whereas rhyming does not (also see Nation and Webb, 2011, on the importance of “generation” as a feature of vocabulary learning activities).
In neither Lindstromberg et al. (2016) nor Snoder and Reynolds (2018) the use of dictogloss was the independent variable of interest, because all treatment groups performed some version of a dictogloss activity. The available research cannot therefore inform teachers whether dictogloss is comparatively effective for the purpose of FS learning, relative to other listening-based activities. The classroom experiment reported further below was meant to help fill this gap in the research by examining the effectiveness of dictogloss for FS learning compared with two other listening-based classroom activities, which most teachers are undoubtedly more familiar with: dictation and answering listening comprehension questions. Apart from measuring the learning outcomes, the study explored factors in the process of performing the dictogloss activity which might influence FS learning. These factors included whether the FSs are correctly reproduced during the text reconstruction activity and, if not, whether they are corrected when students compare their version of the text with the original text. In addition to gauging the effects on FS learning, the study evaluated the extent to which the three listening-based activities affect learners’ retention of text content.

**Chapter 3 Rationale for the Research Design**

There are several reasons for comparing the effects of dictogloss to (a) standard dictation and (b) answering open comprehension questions. One reason is that these share the same modality of input. Course designers and teachers may well appreciate the importance of listening practice but may not always realize that the same aural input can be used in other ways than they are already familiar with. Both standard dictation and answering listening comprehension questions are common classroom activities (at least in the Chinese EFL context), and so their use as comparison treatments is ecologically valid. Researchers have claimed that dictation can support students in processing language at the phrase level, and so it could be expected to be
beneficial for FS learning as well (Oller & Streiff, 1975). The main difference between the

dictogloss activity and dictation activity, however, is the students’ cognitive engagement in the

sense that dictogloss requires more retrieval effort, because it requires students to reconstruct the
text after listening to it, whereas in a standard dictation students write the text almost
concurrently with listening to it. Another difference is that dictogloss presumably invites greater
attention to text content than a standard dictation does. Without adequate comprehension of the
text, it will be impossible to reconstruct it. In a dictation, however, students are probably more
focused on issues of orthography, where it is mostly comprehension of “local” content which
assists with orthographic choices (e.g., that the word must be piece, not peace; their, not there;
etc.). As regards the amount of attention given to text content, dictogloss may therefore be
situated on a cline between dictation (mostly form-focused) and answering comprehension
questions (mostly meaning-focused).

In other regards the three conditions compared in the experiment are similar. First, all
three are listening-based activities. Second, they all involve writing (bi-modal processing). This
holds true also for the listening comprehension condition because questions are open questions
prompting written sentences. Third, the way all three activities were implemented in the
experiment involved pair work and learner-autonomous correction of the work (comparisons
with the original input text). Fourth, the classroom activities involved the same total time on task
(49 minutes).
Chapter 4 Research Questions

The study seeks answers to the following research questions.

Principal Question:

1. Do dictogloss, dictation, and answering listening comprehension questions promote productive knowledge of FSs from an input text to different degrees? If so, which of the activities is comparatively effective?

Given the background discussed above, the hypothesis is that dictogloss will be the most effective of the three activities for retaining FSs from an input text. I expect answering listening comprehension questions to be the least effective for this purpose, owing to its content-focused rather than language-focused nature. Because a standard dictation is more language focused, I expect it to lead to better retention of FSs than answering comprehension questions. However, because it does not require retrieval of the FSs from memory (in the sense of recalling how things were worded in the original text), I do not expect dictation to be as effective as dictogloss.

Because of our special interests in the potential merits of the dictogloss activity, the following secondary research questions focus on aspects of this activity in particular and their association with FS retention.

Secondary Questions:

2. Does successful retrieval of FSs during the dictogloss activity matter for later recall of these MWEs? Put differently, is learners’ successful recall of FSs in the posttests associated especially with their successful retrieval of these MWEs during the text reconstruction activity, or is it associated more strongly with error correction at the stage of the activity where the learners compared their reconstructed version of the text to the original?
This question has both theoretical and practical relevance. At the theoretical level, it has been proposed that it is effortful but successful retrieval that is the most beneficial for retention of knowledge (e.g., Bjork, 1994). If so, we would expect successful retrieval of FSs during the text reconstruction activity to be particularly beneficial. On the other hand, one of the tenets of Swain’s Output Hypothesis is that pushed output makes learners aware of the lacunae in their knowledge, and that their interest in filling those lacunae leads to learning (e.g., Uggen, 2012). If so, then it is perhaps the FSs which a learner failed to reproduce well in the text reconstruction activity that stand a good chance of being remembered after checking against the original input text. At a practical level, teachers who wish to try a dictogloss activity may want to know if it is best to make the activity easy enough for their students to reproduce the text mostly accurately (thus benefiting from successful retrieval) or to leave considerable room for mistakes (thus counting on the benefits of noticing and filling knowledge gaps).

An additional secondary research question concerns learners’ retention of the content of the text:

3. When it comes to retention of text content, is there a difference between dictogloss, dictation, and answering listening comprehension questions?

Since answering listening comprehension questions is clearly a content-focused activity, I expect this to lead to better content retention than the dictation activity, where learners are likely to allocate a lot of attention to formal (orthographic) characteristics. As argued above, dictogloss requires a focus on text content as well, and so perhaps it will also lead to better retention of that content in comparison to a standard dictation.
Chapter 5 Method

5.1 Research Design

This quasi-experimental study included a vocabulary levels test, a pre-test, an immediate and a delayed post-test, and a tally of accurate target FSs on the students’ dictogloss worksheets. Figure 1 shows the relationships between these components.

I also recorded the students’ discussions during pair work as they performed the dictogloss activity and interviewed a few students about their reactions to dictogloss. For lack of space, however, these qualitative data will not be included in this thesis.

![Diagram showing the research design components]

Figure 1. Different components of the research design
5.2 Participants

A total of 142 Chinese EFL first-year university students was recruited for this study. The participants were non-English majors between the ages of 17 to 20. Chinese EFL learners begin to learn English in grade three (8 or 9 years old). Therefore, the participants had approximately 10 years of English learning experience. They were drawn from five classes that were taking the same English courses in parallel. Consent to use their data was obtained from all participants. Volunteers from two of the classes were assigned to the dictogloss condition \((n=56)\), volunteers from another two classes were assigned to the dictation condition \((n=58)\), and the remaining class was assigned to the listening comprehension condition \((n=28)\). Five participants did not take the delayed post-test, thus slightly reducing the post-test samples (dictogloss \(n=55\); dictation \(n=58\); listening comprehension \(n=24\)).

As mentioned, the assignment to treatment condition was done by intact classes. This quasi-random assignment of participants was preferred over random assignment for logistic reasons and because the classroom activities during the respective treatments involved pair work and so it was desirable that the students were used to working with each other.

The paper and pen format of the Vocabulary Levels Test (VLT), updated version by Webb, Sasao and Balance (2017), was used to estimate the participants’ vocabulary knowledge and to help ascertain that most of the words used in the listening text were likely to be familiar to the participants. The VLT includes five word-frequency levels (with 30 test items per level); however, this study only administered the first four levels. A sample of the updated VLT can be found in Appendix D. Schmitt et al. (2001) proposed that a score of at least 26/30 on a given level of the VLT indicates mastery of that level. On the first two levels together, the participants
achieved approximately 56/60, which suggests that they had knowledge of most of the 2,000 highest-frequency word families in English. Table 1 shows the means and standard deviations of the VLT scores.

Table 1

*Means and standard deviations of the participants’ scores per word frequency level*

<table>
<thead>
<tr>
<th>Group</th>
<th>1000 M</th>
<th>1000 SD</th>
<th>2000 M</th>
<th>2000 SD</th>
<th>3000 M</th>
<th>3000 SD</th>
<th>4000 M</th>
<th>4000 SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dictogloss</td>
<td>29.75</td>
<td>0.61</td>
<td>26.86</td>
<td>2.56</td>
<td>22.02</td>
<td>4.21</td>
<td>17.29</td>
<td>4.82</td>
</tr>
<tr>
<td>(n=56)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dictation</td>
<td>29.62</td>
<td>0.85</td>
<td>26.19</td>
<td>2.93</td>
<td>20.36</td>
<td>5.21</td>
<td>16.91</td>
<td>4.66</td>
</tr>
<tr>
<td>(n=58)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Listening</td>
<td>29.71</td>
<td>0.76</td>
<td>26.29</td>
<td>2.93</td>
<td>20.07</td>
<td>4.40</td>
<td>15.07</td>
<td>4.54</td>
</tr>
<tr>
<td>Comprehension</td>
<td>(n=28)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>29.69</td>
<td>0.75</td>
<td>26.47</td>
<td>2.79</td>
<td>20.96</td>
<td>4.73</td>
<td>16.70</td>
<td>4.74</td>
</tr>
<tr>
<td>(N=142)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note.* Maximum score per frequency level = 30.

A pre-test was administered to gauge learners’ prior knowledge of the target FSs (N=10) occurring in the input text (see further below for details) for the three treatments (Appendix E). The descriptive statistics for the scores on this pre-test are shown in Table 2. On average, the
participants showed knowledge of about 2.5 of the 10 target phrases. ANOVA revealed no statistically significant differences between the three groups’ pretest scores ($F=0.74$, $p=.478$).

Table 2

*Means and Standard Deviations for Pretest*

<table>
<thead>
<tr>
<th></th>
<th>Pretest</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$n$</td>
</tr>
<tr>
<td>Dictogloss group</td>
<td>56</td>
</tr>
<tr>
<td>Dictation group</td>
<td>58</td>
</tr>
<tr>
<td>Listening Comprehension group</td>
<td>28</td>
</tr>
</tbody>
</table>

*Note.* Maximum score = 10.

5.3 The Input Text

The same text was used in all three groups. The authentic text (see Appendix F) consists of 183 running words was slightly adapted from the article *Are you a supertaster? Here’s why you should know — and how to find out* ([https://www.cbc.ca/natureofthings/m_features/are-you-a-supertaster-heres-why-you-should-know-and-how-to-find-out](https://www.cbc.ca/natureofthings/m_features/are-you-a-supertaster-heres-why-you-should-know-and-how-to-find-out)). According to Cobb’s (n.d.) Lexical Tutor software ([https://www.lextutor.ca/vp/comp/](https://www.lextutor.ca/vp/comp/)), 85.1% of the running words in this text belong to the 1,000 most frequent word families of English, 92.3% coverage is reached when words belonging to K2 are included, and 95.9% is reached with words from K3. The Vocabulary Levels Test results showed that the participants had mastered at least the 2,000 most frequent words of English, suggesting these students were able to understand at least 92.3% of
the words in the text. To further support students’ text comprehension, a few of the lower-frequency words were pre-taught (see below). In addition, a warm-up activity was implemented to familiarize the students with the topic of the text.

Three recordings were made of an English native speaker reading the text: (a) an uninterrupted version (one minute 11 seconds), a version with pauses between short segments (nine minutes 57 seconds), and a version with pauses between longer segments (typically full sentences) (two minutes). The versions with pauses were prepared for the dictation group. In the latter procedure, the students first listened to the uninterrupted version, and then listened to the recordings with pauses. The recording with pauses between short segments created time for the students to write down what they heard. The recording with pauses between longer-segments gave them time to double check what they had written. The latter procedure matches the usual practice of dictation in Chinese classrooms. With the listening comprehension group, only the uninterrupted recording was used. This was also the version used with the dictogloss group, but to enable the students to write down enough keywords, the recording with pauses between the longer segments was used as well (see below).

The text was divided into two roughly equal parts. Each group listened to the first part according to the procedures outlined further below and completed work on it, before tackling the second part of the text. Dealing with the two parts of the text sequentially rather than asking the students to take in the complete text in one go was intended to make the tasks more manageable.

5.4 Target Items

To measure the FS learning gains resulting from the treatments, thirteen target items had initially been selected from the input text (see Table 3). However, eventually only ten of these
items were included in the data analysis. To be specific, we excluded the first item in the pre-test \((\text{sensitive to})\) because this one was eventually used to illustrate to the students what was expected of them in the test. Another two items \((\text{make a face; on the other hand})\) were excluded because the pre-test revealed that over 70\% of the participants were already familiar with them.

In order to determine whether the selected phrases qualified as FSs, I checked their inclusion in three online monolingual dictionaries of English: Macmillan Dictionary (http://www.macmillandictionary.com), Oxford Learners Dictionary (https://www.oxfordlearnersdictionaries.com) and Longman Dictionary (http://www.ldoceonline.com/dictionary). I also looked up the target items’ frequency in the Corpus of Contemporary American English (COCA) (https://corpus.byu.edu/coca) (Davies, n.d.). All 13 selected phrases are found in the three dictionaries above, and the frequency of all 13 phrases is higher than 200 in COCA. It is therefore safe to say they are considered highly conventional expressions.

Based on my language learning and teaching experience with similar-profile students, I considered the target FSs unlikely to be familiar to the participants in my study, but, as already mentioned, it turned out (from the pre-test) that two of the FSs were already known, and so these were excluded from the analyses.

Due to the authenticity of the input text, the target FSs are of diverse kinds. They are all relatively transparent in meaning, however. The learning challenge was thus not so much comprehension (as in the case of idioms) but remembering the precise composition of the expressions (necessary for accurate production).
Table 3

Target items

<table>
<thead>
<tr>
<th>FS type</th>
<th>Target items</th>
<th>FS type</th>
<th>Target items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adjective + Noun</td>
<td><em>increased risk</em></td>
<td>Other conventional</td>
<td><em>when it comes to</em></td>
</tr>
<tr>
<td>collocation</td>
<td></td>
<td>expression</td>
<td><em>chances are</em></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><em>too...for someone’s liking</em></td>
</tr>
<tr>
<td>Adjective+</td>
<td><em>responsive to</em></td>
<td>Verb phrases</td>
<td><em>lead to</em></td>
</tr>
<tr>
<td>Preposition</td>
<td><em>overwhelmed by</em></td>
<td></td>
<td><em>tend to</em></td>
</tr>
<tr>
<td></td>
<td><em>obsessed with</em></td>
<td></td>
<td><em>have an impact on</em></td>
</tr>
</tbody>
</table>

5.5 Pre-teaching and warm-up

Although the input text appeared suitable for the participants both in content and vocabulary load, it is recommended (Nabei, 1996; Kowal & Swain, 1994; Wajnryb, 1990) to prepare the students for the topic and for potentially unfamiliar lexis that could impede adequate text comprehension. As already mentioned, the original text was divided into two parts, and before playing each part, a warm-up question was presented on a PPT slide (See Appendix E) to prompt discussion about the subject of the text. Six words (three for each of the two parts of the text) were pre-taught (see Table 3). These words and their Chinese translations were presented on PPT slides, and the researcher went over this information and answered any questions from the students. It is worth mentioning that two of the pre-taught words (*obsessed* and...
overwhelmed) were constituents of target FSs. Correct recall of these words in the post-tests may therefore be attributed to the explicit pre-teaching rather than the listening activities, and so it is the ability to supply the other constituent of the phrase (i.e., the preposition) in the post-test that matters for the between-group comparisons. The warm-up and pre-teaching steps were the same for the three treatment conditions.

Table 4

*Pre-taught words*

<table>
<thead>
<tr>
<th>Words</th>
<th>Translation</th>
<th>Frequency</th>
<th>Words</th>
<th>Translation</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Respect</td>
<td>方面</td>
<td>2000</td>
<td>Delight</td>
<td>高兴</td>
<td>2000</td>
</tr>
<tr>
<td>Obsessed</td>
<td>着迷的</td>
<td>4000</td>
<td>Fiber</td>
<td>纤维</td>
<td>3000</td>
</tr>
<tr>
<td>Overwhelmed</td>
<td>不知所措的</td>
<td>3000</td>
<td>Bonus</td>
<td>意外惊喜</td>
<td>4000</td>
</tr>
</tbody>
</table>

5.6 Procedures

Table 5 outlines the general data collection procedures, and Table 6 details the treatments. The experiment was presented to the participants as an extracurricular opportunity for them to improve their English, in addition to their regular classes. The Letter of Information and the Consent Form used for the recruitment of the participants can be found in Appendix B and Appendix C.
### Data Collection

<table>
<thead>
<tr>
<th>Session</th>
<th>Stage</th>
<th>Time</th>
<th>Instrument</th>
<th>Procedure</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>(1) Letter of Information</td>
<td>The Letter of Information and the Consent Form were handed to the potential participants. Only the participants who volunteered to participate in the study proceeded to the following stages, and they were quasi-randomly assigned to one of three groups.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(2) Consent form</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>25</td>
<td></td>
<td>(3) The Vocabulary Levels Test (Appendix D)</td>
<td>Participants in all three groups were asked to sit a pen and paper version of the Vocabulary Levels Test and a Pre-test of the target FSs.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(4) Pre-test (Appendix E)</td>
<td>The Vocabulary Levels Test took 15 minutes to complete. The pre-test of target FSs took 10 minutes.</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>49</td>
<td>(5) Treatments: Dictogloss, Dictation, and Listening Comprehension</td>
<td>Different groups were assigned to different treatments. Each treatment took the same total amount of time. For detailed descriptions of the treatments, see Table 2.</td>
</tr>
</tbody>
</table>
Participants were given the immediate post-test after the treatment. The immediate post-test took 10 minutes to finish.

Participants took the delayed post-test one week after the treatment. It was the same as the immediate post-test.

Table 6
Treatments

<table>
<thead>
<tr>
<th>Stage in Treatment Design</th>
<th>Dictogloss Time (Minutes)</th>
<th>Dictation Time (Minutes)</th>
<th>Listening Comprehension Time (Minutes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a)</td>
<td>Investigator explained the procedures of dictogloss activity and answered relevant questions</td>
<td>5</td>
<td>Investigator explained the procedures of dictation activity and answered relevant questions</td>
</tr>
<tr>
<td></td>
<td>Investigator pre-taught</td>
<td>10</td>
<td>Investigator pre-taught</td>
</tr>
<tr>
<td>---</td>
<td>------------------------</td>
<td>----</td>
<td>------------------------</td>
</tr>
<tr>
<td>(b)</td>
<td>six words and conducted</td>
<td></td>
<td>six words and</td>
</tr>
<tr>
<td></td>
<td>a classroom discussion</td>
<td></td>
<td>conducted a classroom</td>
</tr>
<tr>
<td></td>
<td>discussion</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(c)</td>
<td>Participants listened to</td>
<td>2</td>
<td>Participants listened</td>
</tr>
<tr>
<td></td>
<td>the normal speed</td>
<td></td>
<td>to the normal speed</td>
</tr>
<tr>
<td></td>
<td>recording of the first</td>
<td></td>
<td>recording of the first-</td>
</tr>
<tr>
<td></td>
<td>part of the text twice</td>
<td></td>
<td>part text twice</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(d)</td>
<td>Participants listened to</td>
<td>1</td>
<td>Participants listened</td>
</tr>
<tr>
<td></td>
<td>the long segments</td>
<td></td>
<td>to the short segments</td>
</tr>
<tr>
<td></td>
<td>version of recording</td>
<td></td>
<td>version of recording</td>
</tr>
<tr>
<td></td>
<td>and took notes in order</td>
<td></td>
<td>and wrote down what</td>
</tr>
<tr>
<td></td>
<td>to finish the text</td>
<td></td>
<td>they heard</td>
</tr>
<tr>
<td></td>
<td>reconstruction activity</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Participants listened to the normal speed recording and took notes to finish the listening comprehension questions. After listening to the recording, they had seven minutes to reconstruct the text individually. Participants consulted in pairs 3.5

Participants listened to the longer segments recording and completed the dictation activity. After listening to the recording, they had five minutes to answer the comprehension questions individually.

Participants reconstructed the text in pairs.
<table>
<thead>
<tr>
<th>(g)</th>
<th>Participants compared</th>
<th>5</th>
<th>Participants</th>
<th>5</th>
<th>Participants compared</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>their text to the original text (see Appendix F) and made corrections using pen in another color</td>
<td>compared their text to the original text (see Appendix F) and made corrections using pen in another color</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(h)</td>
<td>Investigator</td>
<td>16.5</td>
<td>Investigator</td>
<td>16.5</td>
<td>Investigator implemented a second classroom discussion and step c to g was redone using the second part of the text.</td>
<td>16.5</td>
</tr>
<tr>
<td>implemented a second classroom discussion and step c to g was redone using the second part of the text.</td>
<td>implemented a second classroom discussion and step c to g was redone using the second part of the text.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
5.7 Instruments

**Pre-test.** The pre-test in this study was a cued recall test, which tested learners’ prior knowledge of the target FSs. Participants were given sentences with blanks for the missing words of the FSs, and with first letters of all the missing words provided to constrain the options. To make sure that students could fill in the blanks if they knew the target FSs, also Chinese translations of the target items were given.

**Listening comprehension questions for the listening comprehension condition.** Participants in the listening comprehension group were asked to answer the open questions in full English sentences, which aimed to provide opportunities for participants in this group to write down the FSs (similarly to the dictation group and the dictogloss group). The comprehension questions are shown in Appendix J.

**Worksheets.** Instructions and pre-taught words were printed on the top of the worksheets. Worksheets for all three groups were separated into two parts because of the arrangement of the activities. For the dictation group, the worksheet was two pages in total to allow students to write down two parts of the text on different pages. For the dictogloss group, each part of the worksheets was two pages (four in total). Participants were instructed to take notes on the first page and to reconstruct the text on the second page if they found that the first page had no more space or was too messy after notetaking. Listening comprehension questions were printed on the listening comprehension group’s worksheets and enough space was given between the questions for the students to write down their answers.

**Immediate and delayed post-test.** The immediate post-test followed shortly after the treatment activities, at the end of the same class. The delayed post-test was administered one week later. There were two parts in the post-tests. The first and principal part tested the students’
recall of the target FSs. The same format as the pre-test (a cued recall test) was used. However, this time only the first letter of the first word of the target FSs was given (instead of giving the first letter of every word). Since the students had now been exposed to the target FSs several times and could rely on their episodic memory of these encounters with the FSs to realize what expressions the test was meant to elicit, it was no longer considered necessary to give as many cues to prevent students from writing potential alternative phrases.

The second part of the post-test assessed participants’ retention of the content of the text by means of three listening comprehension questions (Appendix H). Only three questions were asked because these needed to be different from the questions already used previously with the listening comprehension group, and this was challenging owing to the short nature of the input text. This small test was included primarily to ascertain that also the students in the dictation and dictogloss activities had engaged with the content of the text, but also to explore potential trade-offs between attention to content and to language features. Answering comprehension questions about a text is a content-focused activity, and so can be expected to result in good retention of content. By comparison, writing down phrases during a dictation requires less engagement with the global content of the text but instead prompts a focus on orthography. If so, text content may not be remembered as well. A text reconstruction activity such as dictogloss is meant to promote engagement with content as well as language, and so it is conceivable that this activity is less susceptible to a trade-off effect.
Chapter 6 Data Processing

Responses in the cued FS tests were scored dichotomously (correct or incorrect). No point was awarded if participants failed to provide a constituent word or provided an incorrect word, such as an incorrect preposition (e.g. *of* instead of *to*). No points were given either when participants made spelling mistakes. This may at first sight seem a questionable decision given that the students engaged primarily with aural input. However, the students were given the opportunity to check their own spelling of words against the print version of the model text after their respective listening activities, and so could be expected to learn the appropriate spelling. As regards conjugation errors, however, a distinction was made between totally fixed phrases, such as *when it comes to*, and phrases with variable conjugation, such as *tend to/ tends to/ tended to*. No point was awarded if a mistake was made in the former case (*when it come to*), but a point was given if a mistake was made in the latter case (on condition that everything else was correct). It was decided to be more lenient in the case of variable conjugational patterns because the test was meant to assess participants’ knowledge of phraseology rather than their mastery of morpho-syntax. Another researcher scored 10% of the test sheets, using the above directions, to check the inter-rater reliability, and total agreement was reached.

As to the FSs written on the worksheets of the dictogloss group before the students compared their version of the text to the original, no points were awarded when the target FS was absent, substituted by another phrase, or had no correct constituent words. One point was awarded for partially accurate FSs, and two points were awarded for fully correct FSs. This coding was done with a view to calculating a Spearman bivariate correlation between the participants’ accuracy of FS use on the worksheets and their accurate recall of FSs in the immediate post-test.
The answers to the text comprehension questions in the second part of the post-tests were scored dichotomously. Participants got a point when they correctly answered a question. Grammatical accuracy was not considered. Inter-rater reliability was checked by providing 10% of the data to another researcher. Agreement was 95%, and the cases of disagreement were discussed until a consensus was reached. After this, I (re-)coded the remaining responses myself.
Chapter 7 Analysis

A series of statistical analyses was run in the SAS environment (2018). To answer the principal question, that is, whether the three treatment conditions lead to different rates of FS retention (as gauged by means of a cued recall test), a multilevel model was used to analyse the test data, because recent literature has indicated its superiority over mixed ANOVA (Seltman, 2014). Besides, as we lost some of our participants in the delayed post-test, a multilevel model has the advantage that it allows computing a model with missing data from some participants. As the post-tests were slightly different (and more challenging) than the pre-test (the pre-test gave the first letter of each word in the FSs, but the post-test only provided the first letter of the first word of the target FSs), the pre-test was not used as a time point in the analysis. Instead, the pre-test scores were included as a covariant.

One of the secondary research questions concerned the association between successful recall of MWEs in the posttests and the learners’ successful retrieval of these MWEs as they tried to reconstruct the text during the dictogloss activity. To answer this question, the successfully recalled FSs in the post-tests were traced back to their appearance on the worksheets.

To answer the question whether text content was retained differently by the dictogloss, dictation, and listening comprehension groups, multilevel regression modelling was conducted in SAS as well.
Chapter 8 Results

Table 7 displays the descriptive statistics of the FS recall tests at the two post-test time points (immediate and one-week delayed). A mixed-effects regression analysis was conducted in SAS to detect any significant trends. As fixed effects, we entered time, treatment, pre-test, and the interaction of time*treatment. The participants’ VLT scores and the pretest scores were strongly correlated ($r = .60, N = 142, p < .0001$), however the participants’ VLT scores did not include because they did not significantly improve the model (but see below). Participants were treated as random effects. The treatment predictor had three levels: dictation, dictogloss, and answering listening comprehension questions, while the time predictor comprised two levels: immediate post-test and delayed post-test. Visual inspection of residual plots did not reveal any obvious deviations from homoscedasticity or normality.

The analysis revealed that all the fixed effects parameters contributed significantly to the model for the cued recall tests performance (see Table 8). However, the two-way interaction of treatment and time was also significant, $F(2, 134) = 4.96, p = .008$, suggesting different impacts of the treatments on the immediate and delayed post-test performance. The next analyses therefore examined the differences between groups’ scores at each of the two time points. As this was an unbalanced design (see Table 7), Tukey-Kramer adjustment was used, and the least square means (slightly different from the descriptive statistics in the table) are reported here.

In the immediate post-test, there was no significant difference in FS recall between the dictation group ($M = 4.62, SE = 0.16$) and the dictogloss group ($M = 4.39, SE = 0.16$), $p = .307$, simultaneous difference 0.23, 95% CI: [-0.21, 0.68]. However, these test scores were significantly higher for the dictation group ($p < .0001$, simultaneous difference 1.3, 95% CI:
FORMULAIC SEQUENCE DEVELOPMENT

[0.76, 1.85]) as well as the dictogloss group \((p = .0002\), simultaneous difference = 1.07, 95% CI: [0.53, 1.62]) in comparison to the listening comprehension group \((M= 3.32, SE= 0.23)\).

Table 7

*Means and Standard Deviations for Productive Knowledge of the Target Formulaic Sequences*

<table>
<thead>
<tr>
<th></th>
<th>Immediate posttest</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>Delayed posttest</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(n)</td>
<td>(M)</td>
<td>(SD)</td>
<td>(n)</td>
<td>(M)</td>
<td>(SD)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dictogloss</td>
<td>56</td>
<td>4.55</td>
<td>2.11</td>
<td>55</td>
<td>3.51</td>
<td>1.53</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dictation</td>
<td>58</td>
<td>4.46</td>
<td>2.34</td>
<td>58</td>
<td>3.16</td>
<td>1.65</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Listening Comprehension</td>
<td>28</td>
<td>3.32</td>
<td>2.07</td>
<td>24</td>
<td>3.04</td>
<td>1.81</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note.* Maximum score is 10.

Table 8

*Tests of Fixed Effects for Posttests*

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Numeric (df)</th>
<th>Denominator (df)</th>
<th>(F)</th>
<th>(p)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time</td>
<td>1</td>
<td>134</td>
<td>57.91</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Treatment</td>
<td>2</td>
<td>134</td>
<td>6.37</td>
<td>.002</td>
</tr>
<tr>
<td>Treatment*Time</td>
<td>2</td>
<td>134</td>
<td>4.96</td>
<td>.008</td>
</tr>
<tr>
<td>Pretest</td>
<td>1</td>
<td>134</td>
<td>323.19</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

Table 9

*Performance on the Posttests: Estimates of Fixed Effects*
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>SE</th>
<th>df</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.14</td>
<td>0.29</td>
<td>138</td>
<td>0.49</td>
<td>.621</td>
</tr>
<tr>
<td>Time</td>
<td>IMM</td>
<td>0.32</td>
<td>0.26</td>
<td>134</td>
<td>1.24</td>
</tr>
<tr>
<td></td>
<td>DE</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment</td>
<td>Dictation</td>
<td>0.35</td>
<td>0.29</td>
<td>134</td>
<td>1.21</td>
</tr>
<tr>
<td></td>
<td>Dictogloss</td>
<td>0.31</td>
<td>0.29</td>
<td>134</td>
<td>1.06</td>
</tr>
<tr>
<td></td>
<td>Listening Com</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment*Time</td>
<td>Dictation*IMM</td>
<td>0.96</td>
<td>0.31</td>
<td>134</td>
<td>3.12</td>
</tr>
<tr>
<td></td>
<td>Dictation*DE</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Dictogloss*IMM</td>
<td>0.77</td>
<td>0.31</td>
<td>134</td>
<td>2.48</td>
</tr>
<tr>
<td></td>
<td>Dictogloss*DE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Listening Com*IMM</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Listening Com*DE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pretest</td>
<td>1.11</td>
<td>0.06</td>
<td>134</td>
<td>17.98</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

*Note. IMM= Immediate posttest, DE=Delayed posttest*

Recall was poorer in the delayed posttest, and the drop was more pronounced in the dictation and dictogloss groups than in the listening comprehension group (see Figure 2). As a result, no statistically significant treatment effect was found anymore (e.g., dictogloss vs. listening comprehension $p = .290, 95\%$ CI: [-0.27, 0.88]). In other words, the benefits for FS learning of the dictation and dictogloss activities were short-lived.
Although the VLT test result was removed from the model, I still wanted to explore the possible relationship between the VLT test result for all three treatments and the post-tests’ scores. Thus, Pearson R correlation was conducted, and the result showed that there was a relatively strong positive correlation between the VLT test score and the FSs accuracy in both the immediate post-test ($r = .56, N = 142, p < .0001$) and the delayed post-test ($r = .53, N = 137, p < .0001$). I also computed the correlation between VLT and post-tests scores for each of the three treatment groups separately. The correlation was slightly stronger for the dictation group (immediate post-test: $r = .58, N = 58, p < .0001$; delayed post-test: $r = .54, N = 58, p < .0001$) than in the dictogloss group (immediate post-test: $r = .54, N = 56, p < .0001$; delayed post-test: $r = .51, N = 55, p < .0001$) and listening comprehension group (immediate post-test: $r = .53, N = 28, p = .004$; delayed post-test: $r = .51, N = 24, p = .01$).
Turning now to the question whether it is especially FSs which were retrieved successfully by the dictogloss group during the text reconstruction activity that stand a relatively good chance of being retained in memory, the tallies of FSs on the worksheets revealed that 24.74% of the FSs were absent or incorrect on the worksheets, 26.67% were partially correct, and 48.63% were correct. Spearman bivariate correlation was used to assess the relationship between the accuracy of FSs on the worksheets and the correct responses in the immediate post-test. There was a very strong correlation between the FS accuracy on the worksheets and the correct responses in the immediate post-test, \( r = .91, N = 56, p < .0001 \). A strong correlation was also found in the delayed post-test, \( r = .85, N = 55, p < .0001 \). It thus appears that students with good FS use on the worksheets are the ones that also have good test scores.

Another question to be addressed is whether the three treatment groups recalled the content of the text equally well. We conducted mixed model regression analysis to determine how the different activities, time points, and pre-test scores affect retention of text content. The initial model included time (immediate post-test and delayed post-test), treatment (dictogloss, dictation, and listening comprehension), pre-test, and the interaction of Time*Treatment; however, the model performance did not find a significant Treatment*Delayed interaction, and as it did not improve the model, it was removed. Thus, in the final model, the fixed effect parameters were time \( (F(1,136) = 101.02, p < .0001) \), treatment \( (F(2,136) = 5.92, p < .0034) \), and pre-test \( (F(1,136) = 35.94, p < .0001) \), and the random effect predictor was participant \( (z=4.78, p < .0001, SE=0.07, CI: [0.25, 0.56]) \). Tukey-Kramer adjustment was used to determine the nature of the effects.

As there was no significant Treatment*Time interaction, I looked at time and treatment separately. Students who did the dictation activity had a significantly lower score \( (M= 1.19, p < .0001) \).
than students who did the dictogloss activity ($M=1.62, SE=0.10$), $p = .003$, simultaneous difference -0.43, 95% CI: [-0.71, -0.15]) and students who answered listening comprehension questions ($M=1.66, SE=0.15$), $p = .008$, simultaneous difference -0.47, 95% CI: [-0.82, -0.13]) in post-tests. The difference between the dictogloss group and the listening comprehension group was marginal and non-significant. The scores on the content recall questions fell rather dramatically between the immediate and the delayed posttest, a difference that was statistically significant (immediate post-test score: $M=1.88, SE=0.08$; delayed post-test score: $M=1.09, SE=0.08$; $p<.0001$, simultaneous difference 0.79, 95% CI: [0.63, 0.94]).
Chapter 9 Discussion

Concerning the main research question about the effectiveness of the three listening-based activities for FSs learning, I found no clear evidence that dictogloss is more effective than dictation in this regard, as there was no statistically significant difference between the two groups’ scores in either the immediate or the delayed post-test. This result does not support my hypothesis, which was based on the literature where proponents of dictogloss have presented a compelling theoretical rationale for this type of activity. Still, both the standard dictation and the dictogloss activity led to better recall of FSs in the immediate post-test than the content-focused condition where students answered comprehension questions about the text. This is in keeping with the expectation that dictation and dictogloss require a greater focus on the language code exhibited in a text than answering content-related questions. That said, the learning gains brought about by dictation and dictogloss seem to have been short-lived, because recall of FSs was poor overall in the one-week delayed post-test, and there was no longer a statistically significant difference with the test score of the listening comprehension group. Moreover, the VLT test emerged as a significant predictor in both the immediate post-test score and delayed post-test score, indicating the higher the VLT test score, the greater the chances of FS being retained in both short-term and long-term memory. I am well aware that the VLT test was not designed to assess students’ vocabulary size. However, it is still worth investigating the correlations to explore if the result was affected by the “Matthew effect” (e.g., Stanovich, 1986). The strong correlations thus support this effect where higher proficiency students learn more than the lower proficiency learners.

It is worth mentioning that even in the immediate post-test, few students recalled more than half of the target FSs accurately. Considering the number of times that the students were
FORMULAIC SEQUENCE DEVELOPMENT

exposed to the input text with these FSs as well as the time invested in the activities altogether, this is not particularly encouraging. It is also worth reiterating that the lexical profile of the text was within the students’ reach, that potentially difficult words were explained to them, and that steps were taken to prepare the students for the topic of the text. Even so, dictogloss can be a challenging exercise because it requires a lot of multi-tasking. Students need to focus on content and take notes at the same time; they are then required to try and remember not only what was said but also how it was said; and if they cannot remember the latter, they need to rely on what knowledge they already have of the target language to fill in the gaps. Gallego (2014) found that lower proficiency learners appreciate dictogloss less than advanced learners and suggests this may be because dictogloss necessitates well developed listening skills and knowledge of the target language that is sufficiently proceduralized to help learners cope with the multi-tasking. Put differently, limits to their processing capacity can make it hard for learners to focus on meaning and on form simultaneously (Skehan, 1998). In this context, it is perhaps not so surprising, then, that only about half of the target FSs were reproduced entirely correctly by the students as they tried to reconstruct the text. This, then, left a substantial amount of FS learning to be done at the stage in the lesson where the students were asked to compare their text version to the original and make amendments.

In comparison, doing a dictation requires less multi-tasking, because learners are given time to write down what they have just heard before moving on to the next segment of the text. From the worksheets, we could clearly tell that participants in the dictation group achieved a much higher accuracy than the participants in the dictogloss group during task performance. This means that there is less risk of students getting the lexical makeup of FSs wrong during the dictation activity—because they simply write down what they hear. If the students managed to
write down the FSs accurately in the early stage of the dictation, then this also means they went over these correct written renderings of the FSs on their dictation worksheets again as the recording was played again for them to check their work. On the downside, the students’ engagement with the FSs (and especially the highly familiar words in them) during dictation may have been rather “shallow” in comparison to dictogloss. Dictation simply requires retrieval of short segments from phonological short-term memory, whereas dictogloss requires retrieval after listening to a whole passage, which is far more effortful. According to the notion of “desirable difficulties” (Bjork, 1994), effortful retrieval generates stronger memories. In short, dictogloss may not have lived up to its full potential here because the students often failed to retrieve the target items during the text reconstruction exercise (i.e., the task was too difficult), while dictation may have brought about similarly unimpressive learning gains because it required little cognitive engagement with the target items (i.e., the task was not difficult enough).

With regard to our secondary research question (does successful retrieval of FSs during the dictogloss activity matter for later recall of these FSs?), the results showed that the better the participants did in the actual dictogloss activity, the better their chances were of getting the FSs right in the immediate post-test. This suggests that successful retrieval during the text reconstruction activity matters. This supports the recommendation in publications about dictogloss that the input text should be matched to learners’ proficiency (e.g. Nabei, 1996; Shak, 2006), and it supports the notion that the activity should be implemented in ways that ensure the students reconstruct the text without too many shortcomings.

The third research question investigated the effectiveness of the three listening-based activities with regards to content retention. I hypothesized that dictogloss engages students more with the content of a text than a regular dictation does. This seems supported by the finding that
the dictation group performed more poorly than the listening comprehension group on the post-test items about text content, while the dictogloss group performed on par with the latter. It needs to be acknowledged here that there were only three test items about text content. So, while the data are in line with the prediction that dictogloss invites more engagement with the content of a text than dictation does, this conclusion must remain very tentative, and this matter requires more research.
Chapter 10 Conclusions and Pedagogical Implications

This research project was motivated by the need to expand language teachers’ arsenal of input-based classroom activities that help foster students’ knowledge of formulaic sequences. A couple of publications (e.g., Lindstromberg et al., 2016; Wood, 2010) have proposed dictogloss as a useful addition for this purpose. Previous investigations of dictogloss for FS learning, such as Lindstromberg et al. (2016), did not include comparisons with other activities, however, and so the effectiveness relative to other options available to teachers remained unclear. I therefore decided to conduct a quasi-experimental study in which I compared FS learning gains from dictogloss and two other activities based on the same listening text: dictation and answering listening comprehension questions. The latter two were chosen as comparison conditions because they involve the same (aural) modality of input and because they are common classroom activities which teachers are most likely already familiar with.

The results of the study suggest that dictogloss—as implemented here—is not more effective than dictation for the purpose of FS uptake, although both dictogloss and dictation are more effective than answering comprehension questions about the text. On the positive side, dictogloss appears to engage students with the content of a text to a degree that is comparable to answering comprehension questions, whereas dictation does not.

There were two main reasons why I hypothesized that dictogloss would be more conducive to FS learning than a standard dictation. One was the effortful nature of retrieval of FSs from episodic memory at the text reconstruction stage of the lesson. The other was the awareness-raising function of pushed output, whereby students may experience a gap in their knowledge and will be interested in filling this gap when they next get the opportunity to do so. In the case of dictogloss, this opportunity arises at the stage when students compare their
reconstructed version of the text with the original text. In order to gauge whether these two pathways for learning have the same impact, I tallied the FSs which the students wrote down accurately on their dictogloss worksheets at the text reconstruction stage (reflecting successful retrieval). According to my analysis, it was the students who were relatively successful at retrieving the FSs during text reconstruction who were also relatively successful at recalling them in the posttest. This finding suggests that for dictogloss to be relatively effective, it needs to be implemented in a way that ensures a good success rate at the text reconstruction stage of the lesson. It is worth recalling that in the present experiment, on average only about half of the target FSs were reproduced entirely correctly during the text reconstruction stage of the lesson. Had the students been more successful, then perhaps the learning gains from the dictogloss activity would have been more substantial. More research is needed to examine this possibility. It is nonetheless safe to say that teachers need to make dictogloss manageable, so students reproduce a lot of the text accurately, leaving only a relatively small number of issues to be fixed as they check their work against the original, model text.

It is important to specify that the results from this study concern the learning of FSs. The results are not necessarily generalizable to other elements or features of language. For example, these students may have been less aware of the phenomenon of formulaic language and of the importance of learning FSs than they were of, for example, word learning and grammar learning. If so, they may not have paid much attention to the FSs included in the text. It is indeed common for teachers and materials writers in the Chinese context to put emphasis on single words and grammar patterns rather than conventional word combinations. This may also help to explain why Lindstromberg et al. (2016) obtained better results when they provided the target FSs on the students’ worksheets, so the relevance of these FSs could hardly be ignored.
Chapter 11 Limitations and Directions for Further Research

Several limitations of this study need to be acknowledged. First, while the experiment was substantial in terms of participant numbers, the number of target items of which knowledge was tested was on the small side. Second, I only used one test format, notably a cued productive knowledge test. It is likely that a receptive knowledge test would have revealed more learning gains, which were not captured by the test I used. The productive knowledge test probably also gave an edge to the students who had produced the FSs in their dictogloss and dictation activities. It is possible that a receptive knowledge test would have shown less of a difference between the dictogloss/dictation groups and the listening comprehension group. Third, I only examined effects of the three treatment conditions on the learning of FSs; potential learning of other language items or patterns was not examined, and neither were any potential benefits for the development of listening skills (although it was interesting to find that dictogloss apparently stimulated content-focused listening more than dictation did). Fourth, while the worksheets of the dictogloss group were analyzed for the presence of the target FSs, a meticulous analysis of this kind should also be done for the worksheets of the other groups, to ascertain that the FSs were written down accurately by the students who did the dictation and to tally the number of target FSs used by the students in the listening comprehension group as they answered the open questions. Finally, this was just a single implementation of dictogloss. As discussed above, implementing it with a greater success rate at the text reconstruction stage could furnish stronger support for its usefulness. For example, showing a print version of the input text first, before using the aural version, could alleviate some of the challenges posed by orthography. It is also worth mentioning that some of the students had never done a dictogloss activity before. Familiarizing the students with this kind of activity could increase its effectiveness as well.
This was but a first endeavor to compare dictogloss with other listening-based activities for promoting FS learning. Much more research on this topic would be welcome. For example, future studies could investigate ways of modifying the dictogloss activity with a view to increasing the chances that students will notice specific items. Lindstromberg et al. (2016) did this by giving the target FSs on the students’ worksheets. While this made retrieval from memory unnecessary, it at least ensured that the students saw the expressions in their accurate form throughout the activity. A possibility might be to give only the content words of certain FSs on the worksheet. This would reduce the challenge of reproducing the entire word string, but still prompt engagement with the function words (e.g., preposition, delexicalized verb) to complete it.

Another possibility is to visually enhance target items in the original text when it is handed out to the students to compare their reconstruction with. As discussed in the literature review, visual enhancement of FSs has been found effective in drawing learners’ attention to such items. It may thus reduce the likelihood that students overlook differences between their rendering of a FS and its accurate use in the original text. More explicit teacher-led feedback at the comparison stage is yet another option. It is typical of empirical studies to try and isolate the effect of a single intervention (or independent variable). Yet, it is often a combination of pedagogical steps or components of a lesson that makes a noticeable difference (e.g., Li, Ellis & Zhu, 2016). Combining tricks of the trade is also what teachers tend to do in real classrooms. Future studies could therefore investigate whether combining dictogloss with certain additional activities fosters better FS learning than other listening-based activities (such as dictation or answering comprehensions questions) in combination with the same additional activities.

A broader curricular approach is to teach students explicitly about the importance of FSs and to combine incidental and semi-incidental learning opportunities with regular intentional FS
learning activities. As learners’ general appreciation of formulaic language grows, they may begin to attend more spontaneously to instances of formulaic language in the texts they are exposed to. Again, it would be interesting to evaluate the merits of dictogloss as part of this broader scheme.
References


Candry, S., Elgort, I., Deconinck, J., & Eyckmans, J. (2017). Word writing vs. meaning


FORMULAIC SEQUENCE DEVELOPMENT


Dobao, P. (2014). Meaning oriented tasks such as dictogloss tend to elicit more lexical than grammatical LREs. Language Teaching Research, 18, 497-520.


Learning and Teaching, 6(3), 513–535.


Kuiken, F., & Vedder, I. (2002). The effect of interaction in acquiring the grammar of a second


Szudarski, P. & Carter, R. (2016). The role of input flood and input enhancement in EFL


Webb, S., Chang, A, C-S. (under review). How does mode of input affect the incidental learning of multiword combinations?


Williams, J., & Evans, J. (1998). What kind of focus and on which forms? In C. Doughty, & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 139-155) Cambridge: Cambridge University Press.


Appendices

Appendix A: Ethics Approval Notice

Date: 2 October 2019
To: Professor Frank Boers
Project ID: 114244

Study Title: Listening-based Activities Promoting Productive Knowledge of Formulaic Sequences
Short Title: Listening-Based Activities

Application Type: NMREB Initial Application
Review Type: Delegated
Full Board Reporting Date: November 1 2019
Date Approval Issued: 02/Oct/2019
REB Approval Expiry Date: 02/Oct/2020

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 must also be obtained prior to the conduct of the study.

Documents Approved:

<table>
<thead>
<tr>
<th>Document Name</th>
<th>Document Type</th>
<th>Document Date</th>
<th>Document Version</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEBRIEFING FORM</td>
<td>Debriefing document</td>
<td>10/Jun/2019</td>
<td></td>
</tr>
<tr>
<td>Delayed post</td>
<td>Paper Survey</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Immediate Post</td>
<td>Paper Survey</td>
<td></td>
<td></td>
</tr>
<tr>
<td>In-Class Recruitment Final</td>
<td>Oral Script</td>
<td>25/Aug/2019</td>
<td>Class</td>
</tr>
<tr>
<td>Interview</td>
<td>Interview Guide</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Letter of Information C</td>
<td>Written Consent/Assent</td>
<td>23/Sep/2019</td>
<td></td>
</tr>
<tr>
<td>Letter of Information Interview C</td>
<td>Written Consent/Assent</td>
<td>23/Sep/2019</td>
<td></td>
</tr>
<tr>
<td>Pre-test</td>
<td>Paper Survey</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Work Sheets</td>
<td>Paper Survey</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Documents Acknowledged:

<table>
<thead>
<tr>
<th>Document Name</th>
<th>Document Type</th>
<th>Document Date</th>
<th>Document Version</th>
</tr>
</thead>
<tbody>
<tr>
<td>Procedure Table</td>
<td>Supplementary Tables/Figures</td>
<td>16/Jun/2019</td>
<td></td>
</tr>
</tbody>
</table>

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 REB registration number IRB-00000941.

Please do not hesitate to contact us if you have any questions.

Sincerely,
Appendix B: Letter of Information-Student

**Project Title:** Listening Based Activities Promoting Language Learning  
Principal Investigator  
Dr. Frank Boers, PhD  
Student Investigator  
Xi Yu, 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 because you are a second language learner. If you agree to participate, it is expected that you will be in this study for two weeks, and it will take three hours of your time in total. You can check the detailed procedures in the table at the end of this document. The aim of this study is to investigate how different listening-based activities support language learning. You will complete the tasks as an extracurricular activity, that is, outside your regular courses.

**The rationale of the study**

Most of the literature focuses on the role of reading to support language learning, but few studies consider the benefits of listening activities. Therefore, this research project intends to examine the benefits of relevant listening activities for language learning. The ultimate goal is to be able to inform language teachers and course designers about the kinds of listening-based classroom activities that are particularly useful.

**The assignment of groups**

If you decide to participate, then you will be randomly assigned to one of three groups, and each group will engage in a specific type of listening activities. 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 have 1 in 3 chance of being placed in any group. Neither you nor the researchers can choose what group you will be in. All three groups will use the same listening text, but the exercises based on the listening text will be slightly different (but will take the same time).

**The procedures of the study**
If you agree to participate, you will be asked to engage in six stages (see the table at the end of the document). The activities will resemble normal classroom experience, but with a focus on listening and pair-work exercises about a listening text, which will take place in a regular classroom at your university. Besides, you may be recorded during the classroom activity, but the investigator will first ask for your permission to do so. The investigator will also invite some of you to have a one-on-one interview to ask about your impressions regarding the listening activities. If you do not want to be recorded or interviewed, you can still participate in the other parts of the research and thus benefit from the exercises.

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 an English language learner. This study will be beneficial for you as it will allow you to practice your English. 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 teachers about effective listening activities which they can use in their classrooms.

The option of leaving the study

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 the process of analyzing the data by then.

Data privacy

All the data collected from you will be kept confidential. The results of the research project will be reported in the student investigator’s dissertation and possibly in journal articles and conference presentations. No names of any individual students will be mentioned in these reports. Note that representatives of Western University’s Non-Medical Research Ethics Board may require access to this study-related records to monitor the conduct of the research.

The data will be stored on a secure server at Western University and will be retained for a maximum of 7 years. Your data may be used for future research purposes (e.g., to answer a new research question). By consenting to participate in this study, you are agreeing that your data can be used beyond the purposes of this present study by either the current or other researchers.
The rights of participants

Your participation in this study is voluntary. You may decide not to take part in this study. Even after you give your consent to participate, you will 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 Xi Yu, 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 Research Ethics at the University of Western Ontario (519)661-3036, 1-884-720-9816, email: ethics@uwo.ca. 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.

Table Study procedures

<table>
<thead>
<tr>
<th>Stages</th>
<th>Time (minutes)</th>
<th>Procedures</th>
<th>Procedures</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>25</td>
<td>You will be provided</td>
<td>- You will receive the relevant information about the research.</td>
<td>A regular classroom at your</td>
</tr>
<tr>
<td></td>
<td></td>
<td>two documents:</td>
<td>- You may or may not agree to participate in this study.</td>
<td>university</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1) Letter of Information</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2) Consent form</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>40</td>
<td>(3) Vocabulary Tests</td>
<td>- You will have 40 minutes to finish two vocabulary tests to estimate how</td>
<td>A regular classroom at your</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>much English vocabulary you already know</td>
<td>university</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>A Two-day Interval</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>40</td>
<td>(4) Listening Exercises</td>
<td>- You will be randomly assigned to a group to do the listening exercises.</td>
<td>A regular classroom at your</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>university</td>
</tr>
<tr>
<td>3</td>
<td>20</td>
<td>(6) Vocabulary test</td>
<td>- You will have 20 minutes to finish this test.</td>
<td>A regular classroom at your</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- The test aims to see what you have learned from the listening exercises</td>
<td>university</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A Two-week Interval</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>20</td>
<td>(7) Delayed survey</td>
<td>- You will have 20 minutes to finish a survey about the activity that you</td>
<td>A regular classroom at your</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>participated in.</td>
<td>university</td>
</tr>
</tbody>
</table>
This letter is yours to keep for your future reference.
Appendix C: Consent Form

Project Title: Listening Based Activities Promoting Language Learning
Principal Investigator: Dr. Frank Boers
Student Investigator: Xi Yu

For participants

- You have read and understood the Letter of Information.
- All relevant questions have been explained satisfactorily by the investigator.
- You will be given a copy of the Letter of Information and this consent once you have signed.
- You have understood that you may be recorded during the classroom activity if you agree to this.

I agree to be audio-recorded in this research.

☐ YES  ☐ NO  

______________________    ______________________  _____________________
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 Investigator                  Signature                    Date (DD-MM-YYYY)

About the results of study:

If you would like to receive a summary of the research findings, please leave your email address here: ______________________________
Appendix D: A sample of the Updated Vocabulary Levels Test

To complete this test, match each definition to the word it defines.

<table>
<thead>
<tr>
<th>Land with water all around it</th>
<th>game</th>
<th>island</th>
<th>mouth</th>
<th>movie</th>
<th>song</th>
<th>yard</th>
</tr>
</thead>
<tbody>
<tr>
<td>Part of your body used for eating and talking</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Piece of music</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Below is how it may look after choosing the answers.

<table>
<thead>
<tr>
<th>Land with water all around it</th>
<th>game</th>
<th>island</th>
<th>mouth</th>
<th>movie</th>
<th>song</th>
<th>yard</th>
</tr>
</thead>
</table>
| Part of your body used for eating and talking |      |        |       |       |      | $$
| Piece of music |     |        |       |       | $$ | $$ |
Please fill the blanks with suitable words. Each blank stands for one word, with its first letter given. Hints and their Chinese meanings are given at the end of the sentences.

1. W_______ i ________c_______ t_______ diet and exercise, we know what to do, but we don’t do what we know. (to identify specific topic; 当提到) (Answer: when it comes to)

2. The flavor of cheese is too strong f_______ my l________. (do not like it; 合...的胃口) (Answer: for someone’s liking)

3. Steve is o ________w________ Drake and collects anything and everything connected with him. (unable to stop thinking about something; 对...痴迷) (Answer: obsessed with)

4. A diet that is high in fat can l________ t ________t obesity. (to cause something to develop; 导致) (Answer: lead to)

5. The increase in the number of young people leaving to work in the cities has h________ a________ i __________ o________ (to affect or influence someone or something; 对...有影响) (Answer: have an impact on)

6. The disease has proved r________ t ________ the new treatment. (react quickly; 响应) (Answer: responsive to)

7. Those who smoke have an i________r________ of heart disease. (a rise in the chance that something bad or dangerous may happen; 增加风险) (Answer: increased risk)

8. Cats t________ t ________ eat more in the winter. (likely to behave in a particular way; 趋向于) (Answer: tend to)
9. Harry read a book about his hometown, and he was overwhelmed by homesickness. *(to cause someone to feel sudden strong emotion that he/she does not know how to deal with it; 使不知所措)* (Answer: overwhelmed by)

10. C_______ a_______ that they will be late anyway. *(it is likely that; 有可能)* (Answer: chances are)
Part 1: Up to one-quarter of the world’s population are supertasters. These are people who are more sensitive to flavors than most of us. They do not stand out in other respects, but, when it comes to food, supertasters are obsessed with what they love, and they hate what they do not like. They are the kind of people who will take great delight in their favorite dish, but who will make a face when something tastes too strong for their liking. But does it really matter?

Part 2: The answer is yes, because, if you are a supertaster, this leads you to eating certain foods, and the foods you eat have an impact on your health. Researchers have found that people who are very responsive to bitter food have an increased risk of cancer because they tend not to eat foods that have a lot of fiber. On the other hand, as supertasters can be overwhelmed by strong tastes, being a supertaster can also be a bonus: chances are you don’t feel the need to add as many spoons of sugar to your cup of tea.
Appendix G: Warm-up discussion

Part 1:
This is a recording about “supertasters”.
Let’s find out what supertasters are!

Part 2:
Now we know what supertasters are, so what?
Are supertasters lucky or unlucky in terms of health?
Discuss this question with your partner.
Appendix H: Immediate Post-test

Part I: Please fill the blanks with English phrases. Each blank stands for more than one word, and the number of words you are expected to write is given. Besides, hints and first letters are given.

1. With diet and exercise, we know what to do, but we don’t do what we know. *(four words; to identify specific topic; 当提到)* (Answer: when it comes to)

2. The flavor of cheese is too strong for someone’s liking. *(four words; do not like it; 合...的胃口)* (Answer: for someone’s liking)

3. Steve is obsessed with Drake and collects anything and everything connected with him. *(two words; unable to stop thinking about something; 对...痴迷)* (Answer: obsessed with)

4. A diet that is high in fat can lead to obesity. *(two words; to cause something to develop; 导致)* (Answer: lead to)

5. The increase in the number of young people leaving to work in the cities has had an impact on *(four words; to affect or influence someone or something; 对...有影响)* (Answer: have an impact on)

6. The disease has proved responsive to the new treatment. *(two words; react quickly; 响应)* (Answer: responsive to)

7. Those who smoke have an increased risk of heart disease. *(two words; a rise in the chance that something bad or dangerous may happen; 增加风险)* (Answer: increased risk)
8. Cats tend to eat more in the winter. (two words; likely to behave in a particular way: 趋向于) (Answer: tend to)

9. Harry read a book about his hometown, and he was overwhelmed by a feeling of homesickness. (two words; to cause someone to feel sudden strong emotion that he/she does not know how to deal with it; 使不知所措) (Answer: overwhelmed by)

10. Chances are that they will be late anyway. (two words; it is likely that; 有可能) (Answer: chances are)

Part 2: Please answer the following comprehension questions by using full sentences.

(1) What does the word “supertaster” mean?

(2) Why are supertasters more likely to have cancer?

(3) What is the advantage of being a supertaster?
Appendix I: Delayed post-test

Part 1: Please fill the blanks with English phrases. Each blank stands for more than one word, and the number of words you are expected to write is given. Besides, hints and first letters are given.

1. W_____________________________ diet and exercise, we know what to do, but we
don’t do what we know. (four words; to identify specific topic; 当提到) (Answer: when it comes to)

2. The flavor of cheese is too strong f__________________. (four words; do not like it; 合... 的胃口) (Answer: for someone’s liking)

3. Steve is o ________________ Drake and collects anything and everything connected
with him. (two words; unable to stop thinking about something; 对... 痴迷) (Answer: obsessed with)

4. A diet that is high in fat can l________________ obesity. (two words; to cause something to develop; 导致) (Answer: lead to)

5. The increase in the number of young people leaving to work in the cities has
h________________ (four words; to affect or influence someone or something; 对... 有影响) (Answer: have an impact on)

6. The disease has proved r_______________ the new treatment. (two words; react quickly; 响应) (Answer: responsive to)

7. Those who smoke have an i_______________ of heart disease. (two words; a rise in the chance that something bad or dangerous may happen; 增加风险) (Answer: increased risk)
8. Cats t____________ eat more in the winter. (two words; likely to behave in a
  particular way; 趋向于) (Answer: tend to)

9. Harry read a book about his hometown, and he was o____________________ a feeling
  of homesickness. (two words; to cause someone to feel sudden strong emotion that he/she
does not know how to deal with it; 使不知所措) (Answer: overwhelmed by)

10. C____________ that they will be late anyway. (two words; it is likely that; 有可能)
    (Answer: chances are)

Part 2: Please answer the following comprehension questions by using full sentences.

(4) What does the word “supertaster” mean?

(5) Why are supertasters more likely to have cancer?

(6) What is the advantage of being a supertaster?
Appendix J: Questions for Listening Comprehension Group

Please use full sentences to answer the following comprehension questions.

(1) What percentage of the world’s population are supertasters?

(2) What is the biggest difference between supertasters and the rest of us?

(3) As to finding pleasure in the food they eat, what do supertasters tend to do?

(4) What are the potential risks of being a supertaster?

(5) What is the benefit of being a supertaster?
Curriculum Vitae

Name: Xi Yu

Post-secondary Education and Degrees:
Tianjin Normal University
Tianjin, China
2013-2017 B.A.

The University of Western Ontario
London, Ontario, Canada
2017-2018 M.Ed.

The University of Western Ontario
London, Ontario, Canada
2018-2020 M.A.

Honours and Awards:
Western Entrance Scholarship
2018-2020

Related Work Experience:
Research Assistant
The University of Western Ontario
London, Ontario, Canada
2018-2020

English Teacher
London, Ontario, Canada
AlphaEDU Academy
2019-Present