The relationship between guessing from context, frequency of exposure, and vocabulary acquisition: An investigation into Vietnamese English learners' incidental learning from reading a graded reader

by Doan Trang Nguyen

British Council's Master's Dissertation Awards 2022

Commendation

MA in English Language Teaching

The relationship between guessing from context, frequency of exposure, and vocabulary acquisition: An investigation into Vietnamese English learners' incidental learning from reading a graded reader

by

Doan Trang Nguyen

N0952412

Dissertation submitted to Nottingham Trent University School of Arts and Humanities,

in partial fulfilment of the requirements of the degree of

Master of Arts in English Language Teaching

5/10/2021

(October/ 2021)

Word count

14,996 words

Table of Contents

LIST OF TABLES

LIST OF FIGURES

COPYRIGHT STATEMENT

ACKNOWLEDGEMENT

This part is dedicated to the wonderful individuals without whom I could not have finished this dissertation.

First, I would like to express my sincere gratitude to my supervisor, Dr. Samuel Barclay, who inspired me to do research into vocabulary acquisition. My research would not have progressed this far without his wholehearted guidance, constructive criticism and keen insight through each stage of the process. I am so lucky to have Sam as my lecturer, personal tutor, and supervisor. I have learned a lot from him, and I realized how much I have grown in one year studying and working with Sam. My deep appreciation also goes to all the lecturers and teaching staff who have accompanied me through this MA adventure.

I would like to thank 38 participants whose cooperation is indispensable for this longitudinal study.

I am also thankful for my dear housemate Mai, my course mates Emma and Dona, my friends of 'Hội chị em NTU' and all my friends in the TESOL course for the fond memories we shared. I would like to thank my friends in Vietnam, especially my BFF Dao Duc Anh and Trevor for their encouragement and invaluable comments on my research paper.

My special thanks go to my beloved 'sister/close friend/roommate' Tran Thuy Tien for her warm-hearted support, both academically and emotionally, during my studies in Nottingham. Thank you for always taking care of me.

I wish to express my gratitude to my parents and my brother for their unconditional love and support. Thanks Mom and Dad for giving me the opportunities and experiences that have made me who I am now. This memorable journey would not have been possible had it not been for them.

Last but of no less importance, massive thanks to my boyfriend, Quan, whose love and mental support enabled me to complete this paper.

ABSTRACT

Given that there is a lack of replication studies examining incidental vocabulary learning through reading, this study aims to replicate the research of Waring and Takaki (2003). The original study was adapted to a Vietnamese context to investigate the rate at which vocabulary was acquired and retained from reading a graded reader. The study also expands the original research by exploring the relationship between guessing from context and lexical acquisition since no studies to date has looked at the impact of such individual difference on learning. 25 target nonwords within five bands of different frequency of exposure (17-21 times to those occurring only once) were chosen. Participants were 38 Vietnamese learners of English of intermediate to advanced proficiency. The subjects took the Guessing from Context Test (GCT) (Sasao 2013) before reading the 400-headword graded reader *A Little Princess*. Vocabulary tests (word-form recognition, multiple-choice recognition, and meaning translation) were then conducted immediately and after delays of one week and a month. The results indicated that participants could learn vocabulary incidentally from reading a graded reader, but learning gains differed depending on the test format. Also, words that appeared more frequently were likely to be learnt than those that appeared less frequently, and the minimum number of encounters needed for considerable learning to take place appeared to be eight meetings. Furthermore, there was no significant correlation between guessing from context and vocabulary learning. The data suggest that a great deal of graded reading is needed to develop knowledge of novel lexical items, and guessing from context may not be as closely related to vocabulary learning as previously thought. Implications for teaching and research are also presented.

CHAPTER I: INTRODUCTION

1. Rationale

After a long period of neglect, vocabulary has received increased attention and has been widely recognized as a critical aspect of developing second language (L2) competency (Schmitt 2010). Lexical knowledge plays a key role in language comprehension and production because no verbal communication is possible without words (Sasao 2013). Despite being central to language learning, vocabulary is only one of numerous aspects of language (e.g., grammar, listening, reading, writing, speaking), all competing for attention. Considering the situation of Vietnam, it appears that most of the class time is spent on studying grammar rules, developing reading skills and improving test scores (Pham 2005). Vocabulary, on the other hand, receives relatively less attention, and just a limited number of words is taught through classroom instructions. However, in order to comprehend 98% of a written text in English without support, L2 learners need to know from 8,000 to 9,000 word-families (Nation 2006). Given the insufficient classroom time and a considerable number of word families in English to learn, deliberate vocabulary learning is simply not enough; thus, students should be introduced to some strategies to develop their vocabulary knowledge outside class.

Among many strategies for autonomous vocabulary learning suggested by different scholars (Nation 2013; Schmitt 1997; Webb and Nation 2017), I am personally interested in guessing from context and using graded readers as a way of encountering the L2 outside the classroom. Concerning the former, from the skill-based approach, guessing meaning from context is seen as an integral part of comprehension (Webb and Nation 2017). If it is a reading strategy, guessing will compensate for inadequate vocabulary knowledge. Notably, when looking at the literature, guessing from context is often regarded as a vocabulary learning strategy, and there seems to be an assumption that learners' capability of guessing from context has a positive impact on lexical acquisition (Sasao 2019; Webb and Nation 2017). Nevertheless, research has yet to test this assumption, motivating the current study, which investigated the relationship between guessing from context and vocabulary acquisition.

With regard to the latter, one seminal study of incidental learning from reading is Waring and Takaki (2003), a study that is seen as significant among studies looking at L2 vocabulary acquisition through extensive reading due to their effective methodology (Nation and Webb 2011; Nation 2013). Specifically, the research was conducted in Japanese context, investigating the rate at which vocabulary was acquired and retained after reading a graded reader and examining the impact of frequency of exposure on vocabulary acquisition. Of the 25 words, the meaning of only one item was retained, and none of the items encountered less than eight times were remembered over the three-month retention interval. Even after 18 encounters, there was only a 10-15% chance that the word's meaning would be recalled after three months. The authors also suggested that it might take over 20 repetitions to acquire a target item. The surprising nature of this finding justifies replication in varied contexts. Also, it should be noted that there is a lack of replication studies in the field; hence, replicating Waring and Takaki (2003) would help validate their findings and fill in the gap.

For the above reasons, I decided to conduct a study on *'The relationship between guessing from context, frequency of exposure, and vocabulary acquisition: An investigation into Vietnamese English learners' incidental learning from reading a graded reader'.* Nonetheless, before discussion of my research is possible, it is important to think about an important question first: What does it mean to know a word?

2. The construct of vocabulary knowledge

To understand vocabulary learning, it is essential to define what it means to know a word. Some may assume a word known if knowledge of the word's form and meaning is demonstrated. However, lexical knowledge involves a great deal more. Many researchers proposed different approaches to conceptualising vocabulary knowledge (Henriksen 1999; Nation 2013; Richards 1976); in the scope of this dissertation, Nation's (2013) framework will be adopted as it is likely to be the most thorough description of lexical knowledge produced to date (Barclay and Schmitt 2019).

FIGURE 1
Dimension of word knowledge (Nation 2013)

			What is involved in knowing a word	
Form	spoken	R	What does the word sound like?	
		P	How is the word pronounced?	
	written	R	What does the word look like?	
		P	How is the word written and spelled?	
	word parts	R	What parts are recognisable in this word?	
		P	What word parts are needed to express the meaning?	
Meaning	form and meaning	R	What meaning does this word form signal?	
		P	What word form can be used to express this meaning?	
	concept and referents	R	What is included in the concept?	
		P	What items can the concept refer to?	
	associations	R	What other words does this make us think of?	
		P	What other words could we use instead of this one?	
Use	grammatical functions	R	In what patterns does the word occur?	
		P	In what patterns must we use this word?	
	collocations	R	What words or types of words occur with this one?	
		P	What words or types of words must we use with this one?	
	constraints on use	R	Where, when, and how often would we expect to meet this word?	
	(register, frequency …)	P	Where, when, and how often can we use this word?	

Note: R = receptive knowledge, P = productive knowledge

It can be seen from FIGURE 1 that knowing a word involves various aspects of word-knowledge, including knowledge of form, meaning and use, with each being broken into receptive and productive mastery. This research study, however, will only focus on the form-meaning link.

3. Organisation of the study

This research consists of five chapters.

CHAPTER II: LITERATURE REVIEW

This chapter conceptualises the framework of this research by discussing the literature on incidental vocabulary learning and one type of incidental learning, specifically, extensive reading. Two factors that impact learning from extensive reading, frequency of exposure and guessing from context, are also explored in more detail.

CHAPTER III: METHODOLOGY

This chapter describes the research methodology, including the participants, the research design as well as the instruments and materials chosen for this study. The data collection procedure and data analysis are also included in this chapter.

CHAPTER IV: FINDINGS AND DISCUSSION

This chapter reports the research findings and discusses the results to investigate the central research questions.

CHAPTER V: CONCLUSION

This chapter provides concluding remarks, including a summary of findings and discussion, limitations of the study, and pedagogical and research implications.

CHAPTER II: LITERATURE REVIEW

1. The importance of learning vocabulary

Despite being neglected in the sixties and seventies (Laufer 1989), vocabulary has been increasingly gaining attention over the last few decades. Researchers, instructors and material designers have become more interested in this area; specifically, in vocabulary acquisition in general (Elgort 2011; Ellis and Heimbach 1997; Laufer 2003; Ludwig 1984; Webb and Nation 2017), in vocabulary testing (Laufer and Goldstein 2004; Read 2000; Read and Nation 1986; Sasao and Webb 2017), and in the place of vocabulary in language comprehension and production (Daneman and Green 1986; Laufer and Hadar 1997; van Zeeland and Schmitt 2013; Webb 2009).

A consensus among researchers is that "lexis is the core or heart of language" (Lewis 1993, p. 89). As Wilkins (1972, p. 111) put it, "… while without grammar very little can be conveyed, without vocabulary nothing can be conveyed". Vocabulary contributes significantly to all language skills, including speaking, writing, reading, and listening, and is, therefore, a key unit of almost every aspect in our lives (Webb and Nation 2017). According to Nation (2006), at least 8,000-9,000 and 6,000-7,000 word-families are needed for unassisted comprehension of written text and spoken text, respectively. Hence, learning vocabulary is fundamental for successful L2 use, and it plays a vital role in forming larger structures, namely sentences, paragraphs, and whole texts or speeches (Read 2000). Inadequate vocabulary can lead learners to failure in establishing and achieving comprehensible communication since they could not use the learnt structures and functions (Rivers and Nunan 1991). Furthermore, as Biber and Conrad (2001) put it, developing vocabulary knowledge provides a strong foundation for the acquisition of other aspects of language, namely, phonology, morphology, pragmatics, and syntax, all of which are central to both first language (L1) and L2 learning.

Many learners admitted that many of their difficulties in both productive and receptive language use arose from an insufficient vocabulary (Nation 1990) and realized that vocabulary was crucial in their language learning. In Leki and Carson (1994), students expressed that knowledge of vocabulary was what they wished to achieve the most when attending an English for Academic Purposes writing course. In a protocol study by Cumming (1990), L2 writers were reported to face word-related problems and tended to devote much attention to vocabulary

while writing. Similarly, L2 readers in Huckin's (1995) research relied heavily on lexical knowledge and revealed that poor lexical proficiency was the largest obstacle they wanted to overcome. In general, both researchers and learners are well aware that vocabulary is an indispensable component of second language acquisition and communicative competence. As the importance of vocabulary is established, this research seeks to make a small contribution to incidental vocabulary learning by focusing on two factors that impact learning from extensive reading, namely, frequency of exposure and guessing from context.

2. Approaches to vocabulary learning

Intentional and incidental learning are two common types of vocabulary learning that have been investigated for many years, and there are several ways in which these terms are defined. From the perspective of experimental psychology, intentional learning involves informing learners about a retention test before they study the lexical items, and students might be given explicit instructions on how to engage with those items (e.g., using flashcards to remember word forms and form-meaning mappings) (Lindstromberg 2019). According to Barcroft (2015, p. 42), intentional vocabulary learning is learning words by "consciously attempting to do so", such as studying a particular set of items and trying to learn words from context when reading a text or while viewing word-picture pairs. Essentially, intentional learning is considered a part of 'language-focused learning', one of the 'four strands' of a well-balanced program for learning a L2 proposed by Nation (2007). Regarding incidental learning, there are two definitions that predominate in the literature. In psychology, incidental learning occurs when learners are not told in advance that a vocabulary test will follow an activity (Hulstijn 2001). Meanwhile, within applied linguistics, it is defined as the learning which accrues as a by-product of a meaning-focused task (Chen and Truscott 2010; Ellis 1999). For instance, when reading or listening to a story, a learner might focus on understanding the message, with no intended purpose of learning new vocabulary. Nevertheless, he or she might gradually acquire some new words by seeing or hearing them several times in context. In this case, any lexical items acquired is considered being incidentally rather than intentionally learned. In the scope of this research, incidental vocabulary learning will receive central attention.

3. Incidental vocabulary learning

3.1 Incidental Learning Hypothesis

In their seminal study on incidental L1 lexical development, Nagy, Herman and Anderson (1985) examined how and to what extent incidental vocabulary learning occurred and formulated the 'Incidental Learning Hypothesis', also known as the 'Default Learning Hypothesis'. They proposed that during the school years, students acquired a large amount of new vocabulary mainly through incidental learning from context during free reading. The key determinant of lexical growth was the volume of experience and interaction students had with written language and reading comprehension ability (ibid.). Through a series of studies, they estimated that children might read up to a million L1 words per year and such volume of input would increase the chance for repeated encounters with words in texts, potentially resulting in incidental vocabulary learning. Although some words can be learnt after a single encounter (Webb and Nation 2017), they are seen as exceptional cases to the general rule that knowledge of words is gained in small increments until students eventually have sufficient knowledge to understand and utilize the words. This hypothesis has stimulated a number of studies with research findings and discussion providing strong support and demonstrating that incidental vocabulary learning through reading fosters lexical development. Research shows that L1 words (Jenkins, Stein and Wysocki 1984; Nagy, Anderson and Herman 1987; Nagy, Herman and Anderson 1985), L2 words (Horst, Cobb and Meara 1998; Pitts, White and Krashen 1989; Webb and Chang 2015a), as well as L2 collocations (Pellicer-Sánchez 2017; Webb, Newton and Chang 2013) can be acquired incidentally through reading.

Apart from the quantity of reading, the richness of context is also believed to be crucial to lexical development (Nagy, Herman and Anderson 1985). According to Webb (2008), contextual information of reading materials can affect whether vocabulary growth occurs or not. This is because not all reading materials are the same. While some can facilitate learning by providing useful, rich information to learners, others might fail to stimulate the learning process due to inadequate, ambiguous information that could mislead readers. Take the following two sentences as an example.

(1) Amanda spilt drinks on a classmate and said 'Sorry', sincerely.

(2) Amanda spilt drinks on a classmate and said 'Sorry', derisively.

In the former, readers might correctly infer the positive meaning of 'sincerely' based on the co-text. In the latter, however, readers might make the same assumption about the word 'derisively' while, in fact, it expressed a disrespectful and mocking manner. As Webb (2008) put it, the meaning of an unknown word can be transparent in some sentences while can also be opaque in others. Sometimes, it can be deceptive, causing readers to infer an incorrect meaning (Beck, McKeown and McCaslin 1983). Generally, it seems that words that appear recurrently in sentences that offer some information about the word's meaning will be acquired before those encountered in misleading or less informative sentences (Webb 2008). Therefore, when choosing materials, it is necessary to look at how much of a reading text can be used to learn lexical items, and context should be taken into account to make an accurate assessment of incidental vocabulary acquisition (ibid.).

With reference to L2 vocabulary acquisition, L2 learners can also acquire new words incidentally through repeated encounters; however, the learning gains in L2 appear to be small even after a large amount of study time. Horst, Cobb and Meara (1998) reported that after six hours of reading a graded reader, participants in their study learnt 4.62 words on average. Rodgers (2013) (see Webb 2019) also found that the learning gains made by EFL students through watching a television show over seven hours were 6.4 words. Similarly, Peters and Webb (2018) investigated vocabulary learning from viewing a one-hour television program and found a gain of 13.95% (3.95 words) in the meaning recognition test and 8.31% (3.97 words) in a test of meaning recall. Pavia, Webb and Faez (2019) revealed that the largest gain (from pre-test to immediate post-test) through listening to a L2 song several times was 1.64 words. Part of the explanations for these relatively small gains lies in the small amount of L2 input that L2 learners receive (especially in the EFL context), limiting the potential for incidental learning to take place (Webb and Nation 2017). Although many extensive reading programs have been implemented, limited resources and insufficient time are two major factors that prevent L2 learners from acquiring the same amount of vocabulary gains as L1 learners. This indicates a limitation in the extent to which incidental learning only though reading can fuel L2 lexical development in the EFL environment. That said, incidental vocabulary learning through reading remains extremely important since (1)

there are too many words to learn solely by means of deliberate vocabulary instruction and (2) there is also too much to learn about each word.

3.2 Expanding on incidental vocabulary learning theory

So far, the Incidental Learning Hypothesis has been widely accepted. Recent research has provided some information to help expand on this existing theory and suggested the following points.

An incremental process

It should be noted that the acquisition of vocabulary proceeds, not in an instantaneous, but in an incremental fashion (Hulstijn 2001). Lexical knowledge is acquired in small increments through repeated encounters with time playing a vital role in this learning process. Should words not be re-encountered for a long time, knowledge of those words might be forgotten. In other words, vocabulary knowledge does not always move forward but rather move back and forth along a continuum depending on the number of encounters over time (Nation and Webb 2011). In agreement with Hulstijn (2001), Nation (2013) stated that vocabulary learning is an incremental process, in which knowledge is developed gradually not only at the vocabulary level, but also in terms of each component of word knowledge. Learning a word is not simply connecting the word's form to its meaning, but a complicated process involving the acquisition of grammatical functions, frequency intuitions and sociolinguistics factors through repeated encounters with that word (ibid.). For each encounter, L2 learners might learn some word knowledge from the context, but the knowledge gained might be considered incomplete and easy to decay. Moreover, knowledge of each component of a word might be obtained to different degrees in each encounter, and having knowledge of one aspect does not guarantee that knowledge of another aspect will be gained as well (Webb 2007). Webb (2019) shared the same opinion claiming that incidental vocabulary knowledge is acquired incrementally and there is no dichotomy between knowing and not knowing a word; instead, words should only be seen as partially known since only a small fraction of knowledge about words can be learned at a time.

Quantity of input

A prerequisite for substantial incidental vocabulary learning to occur is that learners receive a large quantity of input (Webb and Nation 2017). Concerning the L1 environment, it is the abundance of input, specifically through repeated encounters with words in context, that facilitates lexical growth. However, when learning L2 in non-English-speaking communities, learners cannot naturally be exposed to a huge amount of L2 input, and this lack of input might limit their incidental vocabulary learning gains (Pavia, Webb and Faez 2019). To foster the growth of vocabulary knowledge for those learners, it is suggested that educators increase L2 input by incorporating programmes, such as extensive reading and extensive viewing of television, in the curriculum because they include sufficient amount of input for repeated encounters with unknown words (Rodgers and Webb 2019). Moreover, as learners can choose what to read or view according to preferences, these sources of authentic input might be enjoyable enough to hold learners' attention for a longer period of time, which also contributes to their incidental vocabulary learning (Day and Bamford 1998).

Varied encounters

Encounters with unknown words variedly and repeatedly is crucial for incidental vocabulary learning to occur as it contributes to all aspects of lexical knowledge (Webb 2019). As discussed earlier, only partial knowledge of a word is acquired through each encounter. Therefore, it is unlikely that knowledge gained through just a few encounters is adequate for learners to fully learn a word. To illustrate, consider the following three sentences that include the word 'take'.

(1) They take their kids to the beach. (= go with)

(2) Skydiving takes a lot of courage. (= need)

(3) The nurse takes his temperature. (= measure)

It can be seen that 'take' conveys different meanings and goes with different collocations in each of the sentences, and this variation cannot be expressed in any one sentence. For most words, there is much to learn about as well. While repetition is essential, varied encounters introducing new information are also of great importance because the degree of difference between prior and current encounters might impact how much about the word can be learned

(Joe 1998). As Hall (1991) claimed, the greater the difference, the stronger the learning. In addition, Webb (2019) stated that through multiple encounters with a word in diverse contexts, learners are likely to learn associations of that word. For instance, if learners meet the word 'football' several times, they might learn other related words, namely, its superordinate association (i.e., sport), its coordinate associations (e.g., basketball, baseball, volleyball), its syntagmatic associates and collocations (e.g., match, play, club). In general, varied encounters with words in input enables learners to develop and consolidate their knowledge of L2 words (Nation 2013).

Spoken versus written input

Research shows that both spoken and written input contribute significantly to incidental vocabulary gains (Pellicer-Sánchez 2016; Peters and Webb 2018; Vidal 2011; Webb and Chang 2012). However, the value of spoken input to vocabulary growth has been underrated (Webb and Nation 2017), which might have occurred due to earlier research's results. In particular, Hayes (1988) claimed that written input provided greater opportunities to encounter unknown words since it contains a higher proportion of low-frequency words. That said, more recent research points out that most people are more likely to encounter L2 spoken input than written input (Kuppens 2010; Lindgren and Muñoz 2013; Peters 2018). Other researchers (Rodgers and Webb 2011; Webb and Rodgers 2009) also indicate that the proportion of low-frequency words found in spoken discourse is not too different from that of written input, and spoken input provides quite similar opportunities for repeated encounters with words. As Webb (2019, p. 231) put it, "the potential for learning vocabulary through spoken input may be at least as great as it is for learning through written input". Webb and Chang (2012) recommended that if these two kinds of input are merged, there will be a greatest chance for incidental vocabulary learning to take place.

Incorporating intentional learning

Research (Nation 2013; Paribakht and Wesche 1997; Zimmerman 1997) has shown that a combination of incidental and intentional learning modes is likely to be superior to the incidental approach alone. It is essential to note that the knowledge gained through incidental vocabulary learning is likely to be receptive knowledge rather than productive. This means learners would be

able to recognize the words' forms (i.e., spoken and written) and understand the meaning when encountering them in input; however, they might not be able to use those words or produce their varying derivations in context (Schmitt and Zimmerman 2002). For those who want to reach productive levels of mastery, it can be difficult to achieve if they only take the incidental learning mode. The reason is that incidental vocabulary learning often provides learners with massive exposure to comprehensible input so that they can 'accidentally' obtain words and does not involve much conscious word manipulations; as a result, it does not generate enough engagement for the words to be acquired fully (Laufer 2005; Nation 2013). In this case, intentional learning should be incorporated in order for a word to be learned at a productive level (Schmitt 2008). Additionally, as Liu and Nation (1985) claimed, learning gains through deliberate learning help set the foundation necessary for incidental learning to occur. In fact, without knowing many of the words in input, learners might hardly acquire words incidentally. Since both incidental and intentional learning have value, these two approaches should be seen as complementary to learning and should be combined to facilitate learners' vocabulary acquisition (Nation 2013).

4. Extensive reading

4.1 Extensive reading and the benefits of reading extensively

Within the incidental learning paradigm, there are different types of incidental learning; one of which is extensive reading. Extensive reading can be defined as an enjoyable reading situation that involves reading large quantities of texts for pleasure, general understanding, or information (Day and Bamford 2002). Students are also encouraged to choose what they want to read for themselves from materials within their level of comprehension (ibid.). According to Krashen's (1985) Input Hypothesis and Affective Filter Hypothesis, learners can develop literacy and acquire L2 only when (1) they understand what they hear or read (i.e., when they obtain comprehensible input) and (2) their affective filters (i.e., the emotion barrier preventing acquisition) are low enough to let the input in. In this sense, extensive reading meets these conditions, especially through graded readers, which are books written within strictly limited vocabulary levels (i.e., any words that are well-beyond learners' current level would be excluded) with accompanying grammatical controls (Nation 2013).

There are two kinds of extensive reading; one focuses on lexical growth, and the other aims at fluency development (ibid.). Regarding the former, the reading texts should contain no less than 1-2% unknown running words to ensure that there are new lexical items to acquire, and no more than 5% (ideally 2%) to make sure that guessing and comprehension can take place. The latter, on the other hand, should contain few or no unknown words because texts having many unknown words would slow down learners' reading; hence, preventing them from achieving the flow and smoothness needed for pleasurable reading. This study, however, focused only on the first type of extensive reading.

Learners can gain a wide range of benefits from extensive reading in terms of language knowledge and general academic success (Pigada and Schmitt 2006). Huckin and Coady (1999, p. 182) described extensive reading as a "pedagogically efficient" approach since it combines reading and vocabulary acquisition. This approach develops learners' autonomy because reading is an individual activity that can be done at any time, either inside or outside the classroom, and therefore, it allows learners of different proficiency levels to learn at their own pace without being 'squeezed' in an inflexible classroom program (Nation 1997). Since learners can read books of their choice which they find interesting, extensive reading can be enjoyable and may motivate learning (ibid.). Gradually, their reading habits and reading skills in general will be improved. Through extensive reading, learners are given the opportunities to gain large quantities of comprehensible input (Nation 2013), particularly in places where there is limited contact with L2. By reading extensively, learners' vocabulary size can be expanded. As suggested by Krashen (1989), learners might encounter more unknown words through extensive reading, bringing chances to infer those words in their context of use and thus might learn their meanings. Similarly, graded readers, provided that they suit learners' level, can contribute significantly to vocabulary learning. As the books are carefully controlled for vocabulary levels with lexical coverage being over 95%, learners can learn the remaining words through dictionary use or guessing from context (Webb and Nation 2017). They also act as a means of establishing repeated encounters with language items learners have previously met. In this way, learners' knowledge of partially-known words will be consolidated and enhanced (Nation and Wang 1999; Nation and Webb 2011).

4.2 Studies of incidental vocabulary learning through extensive reading

There has been a great deal of research examining incidental vocabulary learning from extensive reading (Day, Omura and Hiramatsu 1991; Dupuy and Krashen 1993; Horst, Cobb and Meara 1998; Pigada and Schmitt 2006; Pitts, White and Krashen 1989; Waring and Takaki 2003; Webb and Chang 2015a, 2015b), with findings demonstrating that vocabulary can be learned through extensive reading. Many early studies found minimal learning gains (Day, Omura and Hiramatsu 1991; Dupuy and Krashen 1993; Pitts, White and Krashen 1989). Nonetheless, such studies were critiqued as generally lacking careful control of the research design (Coady 1997; Nation 2013). Notably, most studies looking at incidental learning gains from reading used only one vocabulary test, primarily a multiple-choice (MC) test (Day, Omura and Hiramatsu 1991; Dupuy and Krashen 1993; Pitts, White and Krashen 1989). Having solely one test means that only one type of word knowledge gains could be examined, which limits researchers to a one-dimensional picture of what is occurring as a consequence of the treatment. Moreover, MC tests are not necessarily the most appropriate tool to measure how much learning has occurred for the following reasons. First, learners can make a guess when taking this type of test, which might impact gain scores. Second, the test only measures prompted meaning-recognition while it should assess what is needed in normal reading–the unprompted meaning-recognition (Waring and Takaki 2003). Finally, constructing MC tests with reliability and validity are notoriously complex (Haladyna and Rodriguez 2013). The choice of test types to assess vocabulary learning gains from reading should be carefully considered because their level of difficulty might considerably influence the measurement of learning that has taken place, thus, affecting the gains that can be displayed (Nation 2013). Many researchers (Nation and Webb 2011; Pigada and Schmitt 2006; Webb 2005, 2008), therefore, suggested that several different tests be utilized when carrying out research of this nature so that the range of learning possibilities can be better represented. For this reason, the research presented in this dissertation employed multiple tests of word knowledge.

According to Nation (2013), not until the study of Waring and Takaki (2003) had been introduced was a much richer picture of vocabulary learning from reading extensively depicted. What makes it particularly significant is the use of effective vocabulary measurement. To be

specific, the authors developed a methodology for measuring small gains by implementing not one but three types of measurements that test similar items. These tests are at different difficulty levels: a word-form recognition test (the easiest), a receptive MC vocabulary test, and a word-meaning recall translation test (the hardest). Having multiple tests of the same words like this is important as it enables the researchers to measure how well each item is known. Through those tests, three different levels of knowledge were demonstrated, thus, revealing a range of vocabulary learning that possibly occurred from reading (Nation and Webb 2011).

Moreover, instead of having the immediate post-test only like previous studies, the delayed post-tests were introduced, which brought great value. Waring and Takaki (2003) argued that there would be a lack of retention data due to the use of only an immediate post-test. When taking the test immediately after reading, participants still have fresh memories of the words and might have higher scores than if they take the test some time later. Thus, immediate post-tests may overestimate long-term learning gains. Having acknowledged this issue, Waring and Takaki included two delayed post-tests (after one week and after three months). According to Nation and Wang (1999), for words to be met again before they were forgotten, learners should read a whole graded reader every one to two weeks on average. Therefore, Waring and Takaki used the one-week delayed post-test to represent the time between initial encounter with a word and possible subsequent encounters supposing learners continued to read. Meanwhile, the three-month delayed test reflects the long-term effect of little or no further reading. Using these two tests would, therefore, provide rich data on the quantity of extensive reading necessary for knowledge reinforcement and the influence of spacing subsequent encounters with the L2 words (Nation and Webb 2011).

The results demonstrated that incidental vocabulary learning from reading occurred with gain scores relying on the test formats; nevertheless, few new words were learnt. Surprisingly, only one of 25 items would be retained after three months, and half of the learned words were soon forgotten even when the words were met more than 18 times. For words that were encountered fewer than eight times, participants could not remember any of them after three months. Thus, Waring and Takaki (2003) proposed that participants need to read a massive number of graded readers in order to develop new vocabulary as very few new words were

retained from reading just one book. Also, the authors suggested that the benefits of extensive reading should not only be evaluated by examining vocabulary gains and retention, but by looking at how it helps enhance and consolidate already-known words.

So far, there has been no replication of Waring and Takaki (2003). Despite being seen as less valuable than original research by some researchers (Porte 2012), replication is essential because it helps with threats to validity by extending generalizability (Burn and Grove 2001). Recently, replication has become a great issue since many findings in various fields (e.g., sociology, psychology) do not hold up when other researchers try to replicate them. As Shanahan (2017) explained, wrong or misleading results occurred might be because researchers made mistakes or might bias the results (intentionally or not). This emphasizes the importance of replication since it can be utilized to "overcome design limitations, increase validity of findings, and bring about correction of error" (Fahs, Morgan and Kalman 2003, p. 67). Having acknowledged the necessity and significance of replication, I would like to replicate Waring and Takaki's (2003) research with the attempt to validate their findings.

4.3 Factors that impact learning from extensive reading

Research has demonstrated that there are various factors affecting the learning of lexical items (Peters 2019). Learning gains can be influenced by (1) word-related factors (either interlexical or intralexical factors) such as cognateness (de Groot and Keijzer 2000), L1 frequency (ibid.), L2 frequency (Laufer 1997), part of speech and word length (Barclay and Pellicer-Sánchez 2021), (2) contextual factors, namely, frequency of occurrence (Horst, Cobb and Meara 1998), presentation of semantically un/related items (Tinkham 1993), as well as by (3) individual factors like learners' vocabulary size (Webb and Chang 2015b) and language aptitude (Li 2016). This study investigates the impact of frequency of exposure and guessing meaning from context on vocabulary learning through extensive reading.

4.3.1 Frequency of exposure

Studies have demonstrated that multiple occurrences of words in written input contribute to incidental vocabulary acquisition (Godfroid et al., 2018; Horst, Cobb and Meara 1998; Pellicer-Sánchez and Schmitt 2010; Webb 2007). The reported minimum number of recurrences for considerable learning gains vary; for example, while Horst, Cobb and Meara (1998) suggested at

least eight repetitions, Rott (1999) concluded that six encounters might be enough. Many other studies, nonetheless, indicated that words occurring ten times or more showed the greatest gains (Nation and Wang 1999; Pellicer-Sánchez and Schmitt 2010; Pigada and Schmitt 2006; Webb 2007). The general idea appears to be that the more frequently unknown words are met, the more likely they can be acquired (Webb 2019).

On the other hand, some researchers argued that although more encounters with words could increase the rate of learning, there is unlikely to be a threshold that guarantees learning (Webb and Nation 2017). Several words can be learned with only one or a small number of encounters, while some cannot be learned after plenty of times, e.g., even after 20 encounters (Brown, Waring and Donkaewbua 2008), or, surprisingly, 96 encounters (Saragi, Nation and Meister 1978). As Webb and Nation (2017, p. 90) put it, any proposed number of encounters will be "more than enough for some words, but not enough for others". One reason explaining why there is no such threshold is that there are other factors (as mentioned above) apart from frequency of exposure impact acquiring words in meaning-focused input, and it is challenging to separate their impacts from that of repetition (ibid.). In other words, learning cannot be ensured just by meeting words a certain number of times.

Overall, there seems to be a controversy among researchers in this area, some proposed a certain number of exposures for incidental vocabulary learning to take place, others says that there is not a threshold beyond which learning is guaranteed. These competing research findings among vocabulary studies urge more research, and this present study is an attempt to verify earlier studies' results, specifically examining whether words of varying frequency of exposures are more likely to be acquired and retained or forgotten. Aside from frequency of exposure, this study also considers an individual factor affecting the learning of lexical items, namely, guessing meaning from context, which will be discussed in the next section.

4.3.2 Guessing meaning from context

Guessing meaning from context takes place when learners read or listen to a meaning-focused input, in which their background knowledge or contextual clues might be used to help them infer the meaning of new lexical items they encounter (Webb and Nation 2017). In general, guessing from context involves the following steps:

"1. Determine the part of speech of the unknown word.

2. Analyse the immediate context to try to determine the meaning of the unknown word.

3. Analyse the wider context to try to determine the meaning of the unknown word.

4. Guess the meaning of the unknown word.

5. Check the guess against the information that was found in the first four steps."

(Clarke and Nation 1980, see Webb and Nation 2017, p. 209)

Guessing from context is claimed to be the most frequent and preferred strategy when learners deal with new words in input (Fraser 1999). Nonetheless, it does not guarantee successful guessing of all unknown words; in fact, learners often fail in deriving the appropriate meaning of unknown words from context (Sasao and Webb 2018). As reported in some studies, the success rate was quite low, ranging from 12% to 33% in Parry (1991) and at 25.6% in Nassaji (2003). An explanation for these low rates might be due to the insufficient information of the context, making it hard for learners to infer the meaning; sometimes, the information provided could be misleading as well. Another explanation might relate to individual differences, specifically learners' language learning experiences (e.g., language proficiency, vocabulary size), and it is likely that those who have rich language learning experiences are better at guessing than others. Liu and Nation (1985) explained that having a large vocabulary size might enable L2 learners to focus more on the unknown words and enhance their ability to guess and learn any of those they encounter. Furthermore, when learners become better at the target language, their guessing skill will also improve (ibid.). Notably, even when the guesses are wrong, learners could still gain some knowledge of the word, namely, the word's spelling, part of speech, its collocation, and the grammatical function by focusing on the unknown word itself and the immediate and wider context in which the unknown word is found (Clarke and Nation 1980).

One point to note is that successful guesses do not always result in learning (Brown, Waring and Donkaewbua 2008; Horst, Cobb and Meara 1998; Waring and Takaki 2003), but this does not mean the strategy lacks value. First, guessing can contribute significantly to word retention (Webb and Nation 2017). Sasao and Webb (2018) supported this claim and explained that guessing is a productive strategy that requires learners to actively process data in input, including hypothesis testing about word meaning (Ellis 1994). As Schouten-van Parreren (1996) stated, encountering words in context provides a cognitive hook for the retention of the words.

Fraser (1999) also suggested that guessing word meanings followed by consulting a dictionary could greatly increase the effectiveness of the guessing from context strategy. Second, it might take quite a long time to develop comprehensive knowledge of words through several encounters with words in context since vocabulary learning is an incremental process; nevertheless, the gains made by using guessing from context could help speed up the process (Webb and Nation 2017). Finally, since thousands of words necessary to understand written and spoken discourse cannot be acquired solely by deliberate learning, guessing meaning from context is one of the most essential learning strategies that enables L2 learners to become efficient autonomous learners of vocabulary outside the classroom (ibid.).

Because of its importance, guessing from context was included as a part of Schmitt's (1997) model of vocabulary strategies, and it was one of the common vocabulary learning strategies proposed by Nation (2013) and Webb and Nation (2017). Guessing from context is valuable but there seems to be an assumption among researchers that if learners are good at guessing from context, they are more likely to acquire lexical knowledge (since it is taught as a vocabulary learning strategy rather than a reading strategy). Laufer (2010) argued that researchers should not take this assumption for granted because successful guessing does not necessarily lead to word acquisition and retention. In alignment with this, Huckin, Haynes and Coady (1993) (see Schmitt 2000) stated that being able to guess a word from context does not mean that it will be retained. A possible reason is that should a word be easy to guess, learners will infer it quickly with minimum amount of mental processing so as to continue reading; such shallow processing, thus, might not guarantee that the word is remembered (ibid.). To date, there is no answer for the question of whether being better at guessing from context results in higher learning gains because (1) very little research has looked at this individual variable, and (2) no research has investigated this relationship. This calls for more research examining the connection between guessing meaning from context and incidental vocabulary acquisition, and the current research aims to do so.

This study

In sum, the review of studies above indicates that although there have been many studies investigating incidental learning from reading, few consider individual differences and how those

individual factors may relate to learning. In fact, no studies to date have examined learners' ability to guess meaning from context, which is also a form of individual difference, and how it influences incidental learning from reading. Therefore, the present research would fill this gap by looking at the relationship between that capability and learning. Moreover, among the studies of L2 vocabulary acquisition through extensive reading, the study by Waring and Takaki (2003) is particularly significant due to the effective methodology they developed to measure small learning gains. So far, there is no replication of this study; therefore, I would like to replicate their study with an attempt to validate their findings, especially examining the effect of frequency of exposure on learning. Although replication can be considered unfavourable to some researchers for lacking originality (Porte 2012), doing this is essential as it helps confirm existing findings as accurate and broadly applicable (Santos 1989). Polit and Hungler (1999, p. 233) believed that through replication, "much greater confidence can be placed in the findings of a study". By validating research, replication studies "build evidence" and "promote use of findings in practice" (Fahs, Morgan and Kalman 2003, p. 67). Since Waring and Takaki examined only Japanese participants, which calls for further investigation of learners from other backgrounds, my research will be adapted to the Vietnamese context.

The following research questions (RQs) are addressed:

1. How many new words are acquired from reading a graded reader and retained over time?
2. To what extent are the participants more likely to learn words that appear frequently in the text than the ones which appear less frequently?
3. Is there a relationship between participants' vocabulary learning proficiency indicated from the Guessing from Context Test and their vocabulary learning gains in the immediate test after reading a graded reader? If so, to what extent?

CHAPTER III: METHODOLOGY

1. Participants

For a study that examines a correlation, Dörnyei (2007) suggested that at least 30 participants should be involved. Therefore, the participants in this correlational research were 38 Vietnamese learners of English as a Foreign Language, aged 18-29.

The sampling method applied in this research was convenience sampling. Regarding convenience samples, Dörnyei (2007, p. 98) claimed that "an important criterion of sample selection is the convenience of the researcher". In other words, subjects are chosen for the purpose of the study if they meet certain practical criteria, such as targeted characteristics, easy accessibility, availability and willingness to engage (ibid.). For my research, the participants were selected online through my social media and network; all of whom volunteered to join the study.

The proficiency levels of the participants varied from intermediate to advanced (from B1 to C2 in the CEFR framework), which was determined by their self-report of their standardized proficiency test results (e.g., IELTS, TOEFL, VSTEP). The participants also took the Vocabulary Levels Test (VLT) version 2 at 2000-word level (Schmitt, Schmitt and Clapham 2001) (see Appendix 1) and all scored 26 or higher out of 30, which indicated their sufficient vocabulary to comprehend running words in the graded reader. Notably, none of them had read the book *A Little Princess* before.

2. The design

This was a longitudinal study that utilized a within-group design, in which data were gathered at three different occasions allowing for an intra-individual comparison over time. In longitudinal studies, it is essential that the researchers know when and how frequently to collect data, which can be informed by issues of practicability and fitness for purpose (Taris 2000). Since the study was a replication of Waring and Takaki (2003), the test periods would be similar to theirs concerning the immediate and the one-week delayed post-test. Nevertheless, instead of having a three-month delayed test as they did, I carried out a one-month delayed post-test because of practical reasons. Additionally, there was no pre-test since nonwords were used in the graded reader and the post-tests.

3. Instruments

3.1 Guessing from Context Test

The Guessing from Context Test (GCT) developed by Sasao (2013) measures how well L2 learners can successfully infer the meanings of unknown words encountered in context. As guessing from context is one of the most common strategies used in autonomous vocabulary learning (Paribakht and Wesche 1999), the GCT is a useful tool for providing diagnostic information on learners' guessing skill (Webb and Nation 2017). The test consists of three sections, assessing three crucial aspects in guessing:

(1) Part of speech (i.e., whether test-takers can recognize the part of speech of unknown words),

(2) Discourse clue (i.e., whether test-takers can identify the contextual clue that might help them guess the word's meaning), and

(3) Meaning (i.e., whether they can choose the correct meaning of the target word).

From the results of Sasao and Webb (2018), GCT was claimed to be a valid and reliable measure of the guessing skill. Furthermore, since it matched the purpose of the current study, this test was utilized (see Appendix 2). However, piloting had shown that there might be a ceiling effect in section 1 of the GCT, meaning that there is not enough variance in this section to elicit a correlation. Therefore, the main study additionally uses section 2 and 3 because it might be that any variance regarding the GCT only comes out in those two sections.

3.2 Post-tests

To measure different types of word knowledge, three tests taken from the original study were used. They were (1) word-form recognition test, (2) meaning-translation test and (3) MC recognition test, which can be found in Appendix 3.

With regard to the first one, subjects were asked to highlight words that they could recognize from the graded reader *A Little Princess*. There were 25 nonwords that they encountered when reading the book and 17 distractors[1] to analyze the level of guessing. Correct answers were given one point. Each of the words that was 'selected in error' (i.e., when

---

[1] These distractors were also taken from Waring and Takaki's (2003) study.

participants highlighted a nonword that did not appear in the book) was given one point as well, which was considered false recognition and was used to investigate how much guessing occurred.

The meaning-translation test was an unprompted recognition test in which participants had to either give the exact meaning for 25 nonwords in their L1–Vietnamese or provide a plausible approximate answer (e.g., a near-synonym). For example, the exact meaning of *smort* in *A Little Princess* is *beautiful* ('xinh đẹp' in Vietnamese). However, if participants wrote *gorgeous* ('lộng lẫy' in Vietnamese), they would still be given credit for partial knowledge of the meaning of the nonwords. Furthermore, to encourage responses, participants were given two chances to answer. As for the marking, one point was awarded for correct answers while half point was given to a word having a similar meaning.

Regarding the MC test, this was a prompted recognition test with four options (one correct answer and three distractors) and participants were required to highlight the words that they thought were nearest to the nonwords' meanings. An 'I do not know' option was added to reduce the effect of guessing. Notably, the four choices were the same part of speech. For instance, the nonword *mear* means *money.* Since *money* is a concrete noun, all the options were concrete nouns. The correct answers on this test were counted as one point.

Moreover, participants took the test in strict order to maintain the incidental nature of the activity. Specifically, the form-recognition test was given first because it required the least amount of word knowledge. Next, to prevent the transfer of knowledge from one test to another, the meaning-by-translation test was given before the MC test. One point to note here is that as I intend to run a correlation analysis between the number of words that have been learnt (i.e., the result of the immediate post-test) and GCT score, it is necessary that there is no floor or ceiling effect in the immediate post-tests. Based on the pilot study findings, no floor or ceiling effects were found in these post-tests, meaning that they could be used in the main study.

4. Material
The graded reader
There are two points that should be borne in mind for successful guessing to occur when conducting research of this nature. First, new vocabulary can only be inferred from context when learners read at a high level of text comprehension and coverage (Hu and Nation 2000; Laufer

and Ravenhorst-Kalovski 2010; Nation 2013). As Hu and Nation (2000) suggested, the optimal rate appeared to be between 96% and 99% coverage of the known running words. Second, for a word to be learnt, it is necessary that L2 learners encounter that word many times (Pellicer-Sánchez 2016; Webb 2007). As mentioned in the literature review, research has not specified a clear number of repetitions for incidental learning to occur since there are other vital factors that affect learning. However, some researchers proposed that the intermediate repetition goal for incidental learning seems to be at least 10 times or more (Pellicer-Sánchez and Schmitt 2010; Saragi, Nation and Meister 1978; Webb 2007).

To achieve the desired coverage rate, the 400 headwords graded reader *A Little Princess* was chosen with test items being integrated into. As all participants in this study were intermediate or above learners of English while the book was designed for A1/A2 level learners, this book should present no great problems lexically and rather easy for them to read. Choosing a simple reading material like this assures that the surrounding co-text for test items is familiar, allowing me to investigate the acquisition rate based on the test items only. The graded reader was thoroughly piloted with a group of four participants whose data were kept as part of the present study. These participants reported that the book was easy to understand, not too long, and they could finish reading in less than an hour at a reasonable reading speed. Another reason for selecting this graded reader is because of its decent spread of test items concerning the occurrence rate criteria, which will be described below.

5. Target words

5.1 Rationale for the use of nonwords

Similar to Waring and Takaki's (2003) design, some adjustments were made to the graded reader, specifically by changing the spelling of 25 test items, henceforth called *nonwords*. The main advantage of using nonwords is that it helps the researcher ensure that (1) no words would be known before reading and (2) no words would be encountered after reading because meeting the words later may affect participants' recall on the delayed post-tests. Since words are symbols of meanings, making a change in the symbol (its spelling) has both construct validity and face validity, supposing it conforms to normal spelling and collocation conventions (Brown, Waring and Donkaewbua 2008). This is because "the unknown word may represent a familiar concept

and so the new label for that familiar concept is being learned" (Nation 2013, p. 358). Moreover, there is no difference in utilizing L2 words and nonwords since the success rate of real words and nonwords training is compatible (de Groot and Keijzer 2000, see Schmitt 2010).

25 target words in the graded reader were replaced by nonwords that looked like plausible English words and took on English spelling conventions. To illustrate, the words *money*, *new*, and *night* were rendered *mear*, *tantic*, and *cadle* respectively in their nonword forms in both the reading and post-tests. To ensure reliability, the nonwords were the same as those in the study of Waring and Takaki (2003) since they had already been tested for plausibility by native speakers of English. Four Vietnamese participants in the pilot study, who were part of the experiment, also confirmed that they did not have any difficulty in pronouncing these nonwords. The nonwords were left unmarked for natural reading, which meant they were not highlighted in any way (e.g., underlined, bold, coloured) in the text. Definitions and glosses were not provided as well.

5.2 Controlling the word-frequency variable

The current study controlled for the word-frequency variable with the intention of determining the number of reoccurrences a word needs to be met in reading for it to be acquired. Accordingly, to answer RQ2, words of differing frequencies of recurrence needed to be chosen. It was also important to decide which types of words should be selected. For this research, adjectives and nouns were selected because they are generally easier to guess than adverbs (Laufer 1997). Regarding verbs, they often appear with their different inflections in various tenses, making it hard to ascertain whether the word is known and how frequently the word has occurred in the text (Brown, Waring and Donkaewbua 2008); thus, verbs were not opted for.

Having observed the recurrences of words in the graded reader via the website Compleat Lexical Tutor (Cobb n.d.), I created five groups of test words with different frequency of exposure. The groups were as followed:

Group 1: five words appeared 17-21 times;

Group 2: five words appeared 13-14 times;

Group 3: five words appeared 8-12 times;

Group 4: five words appeared 4-5 times;

and Group 5: five words appeared only once in the book.

The sum of the number of occurrences of 25 items in the five categories that need to be acquired was 237. Because *A Little Princess* has 5925 words in text (602 types), this made for 96% coverage of the running words (95.84% coverage by types), which met the criteria of 96-99% set earlier for successful guessing from context. Thus, it could be assumed that the participants would know all the remaining other words in the graded reader as the reading level (Level 1- A1/A2) is below their ability.

It is worth noting that although the list of L2 words and their nonwords equivalent was adopted from Waring and Takaki's (2003) study, some adaptations were made to overcome the original's limitation. First, a limitation that the authors stated in their work is the choice of two words *years* and *yes*. They were reported to be comparatively easier to guess than other words, which compromised the experimental data in some way and were removed from the data analysis. As a result, I decided to substitute these words. With the help of AntConc 3.5.8, a program that allows users to investigate the word list, the frequency and the concordance, two new words (*servant/s* and *child/children)* were selected due to their word type (noun) and occurrences in the book. Second, Waring and Takaki (2003) only made changes to words by adding plural morpheme -s to nouns. Meanwhile, in the present study, all forms of a word were modified, including adding inflectional (e.g., -s) and derivational suffixes (e.g., -ly) to words as well as creating the superlative and comparative forms of the adjectives. Please see Appendix 4 for the list of L2 words, nonwords equivalent and the number of recurrences in the graded reader.

6. Data collection procedure

First, subjects were required to take the GCT, and their tests were marked afterwards. Next, they had one hour to read the graded reader *A Little Princess* online without note-taking or using any other supplementary materials (e.g., dictionaries, search tools). They were told that there would be a test after reading; however, no detailed information was given to prevent them from focusing on the test's content while reading. Once finishing the text, subjects did the immediate post-tests in strict order as explained earlier: the word-form recognition test, the meaning-translation test and the MC recognition test. The graded reader was uploaded to Google Drive and shared with participants in view-mode (read-only), meaning that they could not download, copy or print the text. When participants finished reading, their access to the file

would be removed to ensure that they did not look at the story while doing the post-tests, and the nonwords could not be met again. This is because the current research aims at measuring subjects' recognition and recall of the target vocabulary items (i.e., the nonwords that had been integrated into the book).

One week later, participants were tested the second time, and one month after the first administration, they retook the test. Both of the delayed test administrations were unannounced. Though the delayed tests were the same as the immediate post-tests, the item order in each test was changed to control for a potential learning effect from those tests. Test results were collected for data analysis after each test administration. Notably, since no changes were made to the instruments and material after the pilot study, meaning that the data collection procedure of the pilot study and the main study was similar, the results of four participants in the pilot study were included in this research.

7. Data analysis

Quantitative Analysis

This research adopted a quantitative approach. The collected data were analyzed using Statistical Package for the Social Sciences (SPSS), in which both descriptive and inferential statistics were presented. Also, the results were illustrated visually using tables and charts. To answer the first two RQs, the mean and standard deviations by test types and by frequency of exposure for three test administrations were calculated and compared. Regarding RQ3, a correlation analysis was conducted using two continuous variables: the GCT scores and the learning gains indicated from the immediate post-tests. The correlation coefficient produced would help answer whether there is a relationship between these two variables. Concerning inferential statistics, this research adopted the analysis of variance (ANOVA) and the Kruskal-Wallis test. A repeated measures ANOVA helped answer if there are significant differences between test administration times on three different levels of knowledge, namely, form recognition, meaning recall, and meaning recognition. Meanwhile, the Kruskal-Wallis test revealed whether learning gains indicated from each test type at the immediate post-test were significantly influenced by frequency of exposure. Kruskal-Wallis test was utilized because it

allows the researcher to compare more than two conditions (Field 2013), in this case, to compare five groups of test words with different occurrence rates using pairwise comparisons.

Chapter summary

This chapter began by presenting information related to the participants and the design of the research. The descriptions and justifications for instruments and material used in this study including the GCT, the post-tests, the graded reader, and the target words were also discussed. Additionally, the data collection procedure was covered, followed by the data analysis. In the next chapter, research findings and discussion will be presented.

CHAPTER IV: FINDINGS AND DISCUSSION

1. Findings

The result of the study will be presented according to three RQs.

1.1 RQ1: How many new words are acquired from reading a graded reader and retained over time?

The mean scores (with standard deviations) by test type over the three test administrations are presented in TABLE 1 and are shown graphically in FIGURE 2. The results will be reported in detail below.

TABLE 1

The mean scores and standard deviations (SD) by test type for three test administrations

Test	Administration 1 (Immediate post-test)		Administration 2 (One week delay)		Administration 3 (One month delay)	
	Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*
Word-form recognition	17.61	(2.77)	14.63	(4.37)	15.58	(4.10)
Meaning-translation	8.67	(4.73)	5.46	(5.05)	5.20	(4.74)
Multiple-choice recognition	14.50	(3.98)	13.08	(4.43)	12.68	(4.34)

Note. Max= 25, n= 38.

FIGURE 2

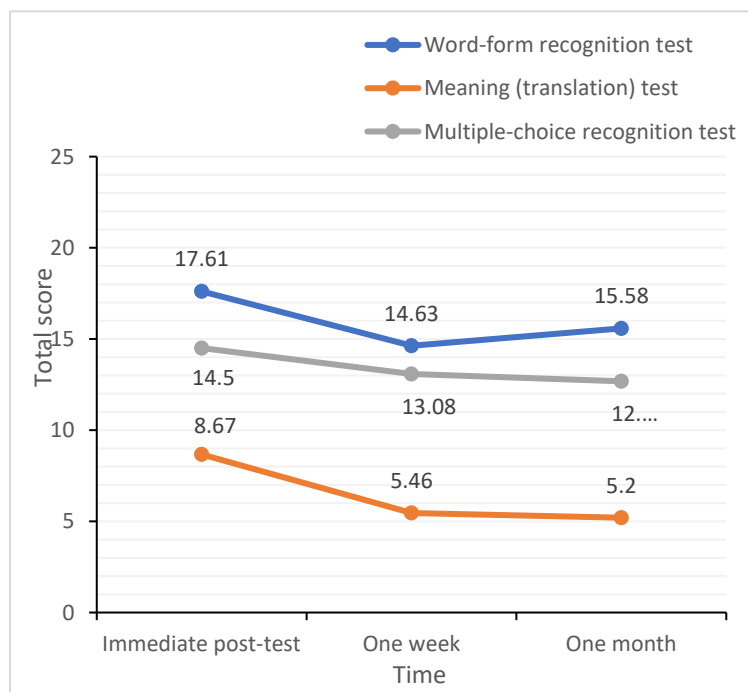*Mean scores by test over the three test administrations*

Table 1 and Figure 2 illustrate that the mean score of 25 items on the immediate word-form recognition post-test was 17.61 (SD= 2.77), which dropped to 14.63 (SD= 4.37) after a week and to 15.58 (SD= 4.10) after a month. Regarding the meaning-translation test, participants were able to learn 8.67 words (SD= 4.73); however, there was a moderate decline in the mean scores of the test items as it dropped to 5.20 (SD= 4.74) after one month. In terms of the multi-choice recognition test, results showed that a mean of 14.50 words (SD = 3.98) were learned over the treatment. That figure decreased to an average of 13.08 (SD = 4.43) and 12.68 (SD = 4.34) on the one week and one month delayed tests respectively.

Overall, subjects learned about 60% of the items on the MC recognition test, and more learning was found on the word form recognition test, with acquisition rate being at 70.44%. Meanwhile, just over a third of the words (34.68%) were learned at the immediate meaning-translation test. There was generally robust retention to the recognition tests, with about 12% of accrued knowledge shown to have decayed over the one-month retention interval. This shows that the treatment produced relatively robust lexical knowledge. In contrast, the meaning-translation test scores decreased more considerably as 40% of the learning gains were lost over a month. As the line chart suggests, the mean scores for the meaning-translation test seem to be much lower than those on the form recognition and meaning recognition test. This is unsurprising given the greater task difficulty of the translation test. In fact, studies have pointed out that meaning recall tests are generally harder than form recognition and meaning recognition tests (González-Fernández and Schmitt 2019; Laufer and Goldstein 2004).

A repeated measures ANOVA revealed that there were statistically significant differences between test administration times and three different levels of knowledge. At the level of word-form recognition, a statistically significant difference was found (F(2, 111) = 5.89, p = .004). Post hoc comparisons showed there was a difference between the immediate test and the one-week delayed test, p = .003. The one-month delayed test did not differ significantly from either the immediate test or one-week delayed one.

At the level of meaning recall, there was a statistically significant difference at the $p < .05$ level for the three time administrations: F(2, 111) = 5.91, p = .004. The difference lay between the immediate test and the first delayed post-test (p = .014) and also between the immediate test

and the second delayed post-test (p = .007). There was no statistically significant difference between the one week and the one month delayed test (p = .970). Finally, there was no statistically significant difference between test administration times on the MC recognition test, F(2, 111) = 1.87, p = .159. Thus, the results showed that in terms of word-form recognition and meaning translation test, time appeared to affect participants' vocabulary retention; however, no impact of time was found on knowledge at the level of meaning recognition.

1.2 RQ2: To what extent are the participants more likely to learn words that appear frequently in the text than the ones which appear less frequently?

1.2.1 Word-form recognition test

TABLE 2 and FIGURE 3 illustrate the mean scores by occurrence rate of the three test administrations of the word-form recognition test.
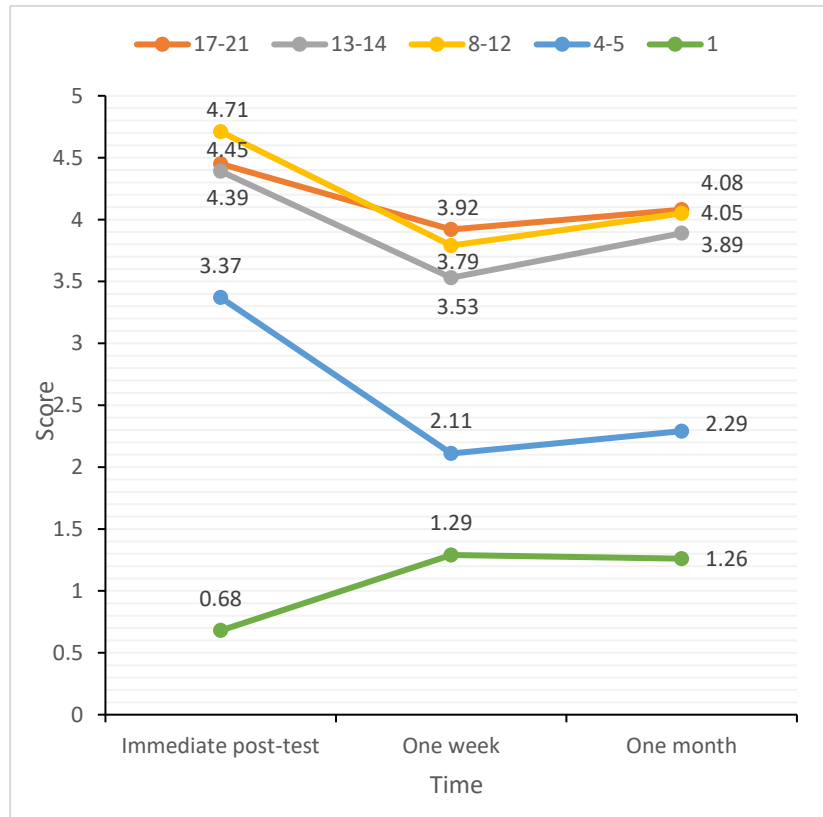
TABLE 2

The mean scores by frequency of exposure for three test administrations on the word-form recognition test

Test	Frequency of exposure				
	17-21	13-14	8-12	4-5	1
Immediate post-test	4.45	4.39	4.71	3.37	0.68
SD	(1.07)	(0.71)	(0.56)	(1.29)	(0.73)
One week	3.92	3.53	3.79	2.11	1.29
SD	(1.04)	(1.16)	(1.15)	(1.52)	(1.12)
One month	4.08	3.89	4.05	2.29	1.26
SD	(0.98)	(1.05)	(1.07)	(1.43)	(1.33)

Note. Max= 5, n= 38.

FIGURE 3

Mean scores by frequency of exposure on the word-form recognition test over one month

As shown in Table 2 and Figure 3, the mean score of words appearing 17-21 times was 4.45 (89%), which slightly declined to 4.08 (81.6%) over a month. For words that appeared only once, the mean scores stayed about the same, at around 1. There was a tendency that the more frequent items, especially words appearing more than eight times, had higher (form) recognition rates than the less frequent items.

A Kruskal-Wallis H test showed that the learning gains indicated from the word-form recognition test at the immediate post-test were significantly affected by frequency of exposure, $H(4) = 114.185$, $p = .000$. Pairwise comparisons with adjusted p-values showed that there were statistically significant differences between numbers of words learnt when participants encountered words appearing only once compared to those appeared 4-5 times ($p = .000$, $r = -0.573$), 8-12 times ($p = .000$, $r = -1.055$), 13-14 times ($p = .000$, $r = -0.894$), or 17-21 times ($p = .000$, $r = -0.985$). There were also statistically significant differences in learning gains between those

seeing words that appeared 4-5 times and those seeing words that appeared 8-12 times (p = .000, r = -0.482) and 17-21 times (p = .003, r = -0.412). However, there were no significant differences in learning gains between encountering words that appear 13-14 times and 4-5 times (p = .052, r = -0.320), 8-12 times (p = 1.000, r = 0.161), and 17-21 times (p = 1.000, r = -0.092). Finally, there were also no significant differences between 8-12 times and 17-21 times (p = 1.000, r = 0.070).

TABLE 3

Mean scores and number of errors on the word-form recognition test for the three test administrations

	Administration 1 (Immediate post-test)		Administration 2 (One week)		Administration 3 (One month)	
	Correct	Selected in error	Correct	Selected in error	Correct	Selected in error
Mean	17.61	1.11	14.63	1.55	15.58	2.03
(SD)	(2.77)	(1.33)	(4.37)	(1.68)	(4.10)	(1.84)
Adjusted means*	16.5		13.08		13.55	

Note. Max= 25, n= 38. *The adjusted means were calculated by subtracting the number of incorrect items from the number of correct ones.

FIGURE 4

The rate of correct and missed recognition on the word-form recognition test over a month
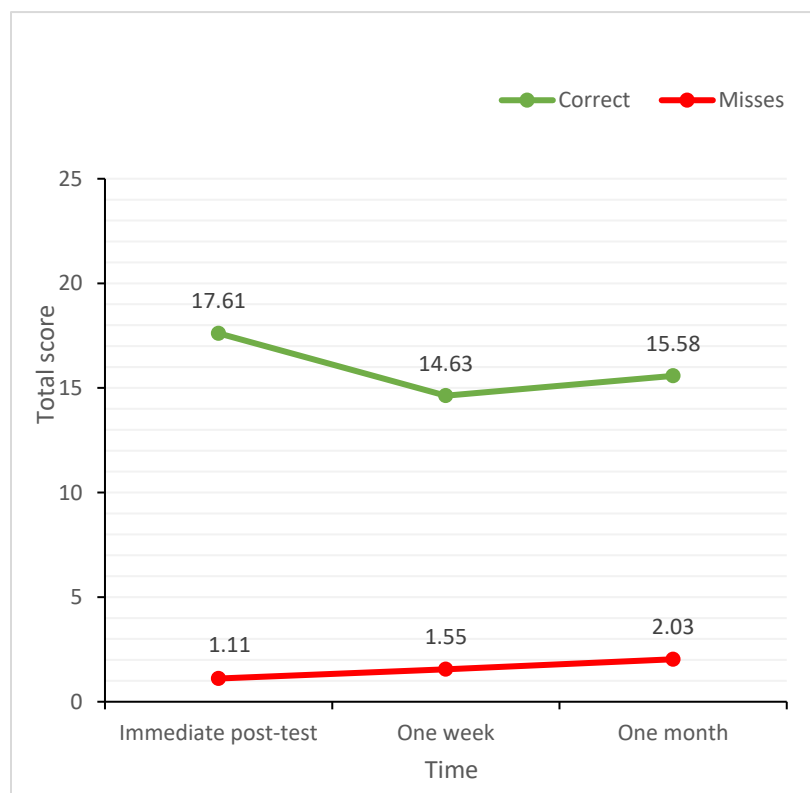
TABLE 3 and FIGURE 4 compare the number of word forms that were 'selected in error' or selected correctly. 'Selected in error' refers to items that were chosen by learners despite not occurring in the graded reader, and is used to indicate how often the participants were guessing. From the results, the mean number of incorrect nonwords chosen gradually inclined from 1.11 items (6.3% of the correct score) on the immediate post-test to 1.55 (10.6%) after a week and to 2.03 (13%) items after a month. After correction, the adjusted mean scores were 16.5 at the immediate post-test, 13.08 after one week, and 13.55 after one month. This seems to suggest that (1) a proportion of vocabulary knowledge had been lost as time passed, and (2) more guessing had occurred over time, in other words, participants were less sure of their knowledge and more willing to guess incorrectly.

1.2.2 Meaning (translation) test

The data by recurrence rate for the three test times on the meaning translation test are demonstrated in TABLE 4 and FIGURE 5.

TABLE 4

The mean scores by frequency of exposure for three test administrations on the meaning translation test

Test	Frequency of exposure				
	17-21	13-14	8-12	4-5	1
Immediate post-test	3.00	2.01	2.59	0.99	0.08
SD	(1.46)	(1.64)	(1.25)	(1.28)	(0.27)
One week	1.97	1.12	1.61	0.74	0.03
SD	(1.61)	(1.45)	(1.41)	(1.29)	(0.16)
One month	1.82	1.04	1.59	0.72	0.03
SD	(1.42)	(1.28)	(1.54)	(1.23)	(0.16)

Note. Max= 5, n= 38.

FIGURE 5

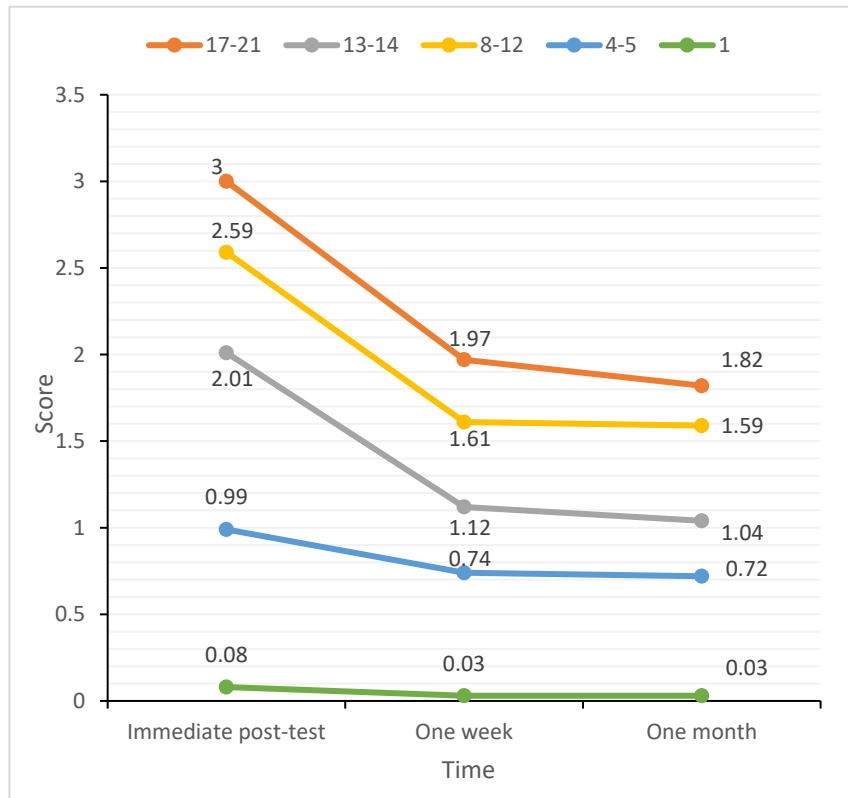*Mean scores by frequency of exposure on the meaning-translation test over one month*

Table 4 and Figure 5 show that participants were able to learn 60% of the words that appeared 17-21 times. Of the five words met 13-14 times, 2.01 of them were learnt. Of the five words met 8-12 times, 2.49 were acquired. Concerning words that appeared less frequently, the acquisition rate seems to be lower: for words that appeared 4-5 times, around one out of five words were learnt; as for words met only once, learning was unlikely to occur.

There was an overall decrease in the mean scores for each of the test administrations. Regarding the immediate post-test, words encountered 17-21 times had a mean score of 3.00 of 5 items (60%), but after a month, the score decreased to 1.82 (36.4%). Words appearing 13-14 times and 8-12 times shared a similar pattern. Concerning the former, the mean score dropped from 2.01 (40.2%) to 1.04 (20.8%), whereas that of the latter fell from 2.59 (51.8%) to 1.59 (31.8%) after one month. Items met 4-5 times registered a score of around 1 (20%) on all tests. As for words that occurred only once, they had a score of near 0 on three test administrations. In

general, words that appeared fewer than eight times seem to be more difficult to recall than those encountered more than eight times.

The number of words learnt indicated from the meaning translation test at the immediate post-test were significantly impacted by frequency of exposure, $H(4) = 88.606$, $p = .000$. Pairwise comparisons showed that there were statistically significant differences in learning gains of items that occurred only once and 8-12 times ($p = .000$, $r = -0.835$), 13-14 times ($p = .000$, $r = -0.642$) and 17-21 times ($p = .000$, $r = -0.922$). Additionally, there were statistically significant differences between learning gains when words appeared 4-5 times compared to those appeared 8-12 times ($p = .000$, $r = -0.514$) or 17-21 times ($p = .003$, $r = -0.601$). Nonetheless, there were no significant differences in learning gains of words that met 13-14 times and 4-5 times ($p = .052$, $r = -0.321$), 8-12 times ($p = .925$, $r = 0.193$), and 17-21 times ($p = .145$, $r = -0.280$). Finally, there were no significant differences between 1 time and 4-5 times ($p = .051$, $r = -0.321$), and between 8-12 times and 17-21 times ($p = 1.000$, $r = -0.088$).

1.2.3 Multiple-choice recognition test

The mean scores by frequency of exposure for the three test times on the MC recognition test are presented in TABLE 5 and FIGURE 6.

TABLE 5

The mean scores by frequency of exposure for three test administrations on the multiple-choice recognition test

Test	Frequency of exposure				
	17-21	13-14	8-12	4-5	1
Immediate post-test	4.11	3.74	3.89	2.45	0.32
SD	(1.05)	(0.96)	(1.19)	(1.80)	(0.46)
One week	3.97	3.18	3.32	2.18	0.42
SD	(1.04)	(1.17)	(1.17)	(1.62)	(0.67)
One month	4.03	2.92	3.39	2.11	0.24
SD	(1.01)	(1.33)	(1.23)	(1.55)	(0.53)

Note. Max= 5, n= 38.

FIGURE 6

Mean scores by frequency of exposure on the multiple-choice recognition test over one month
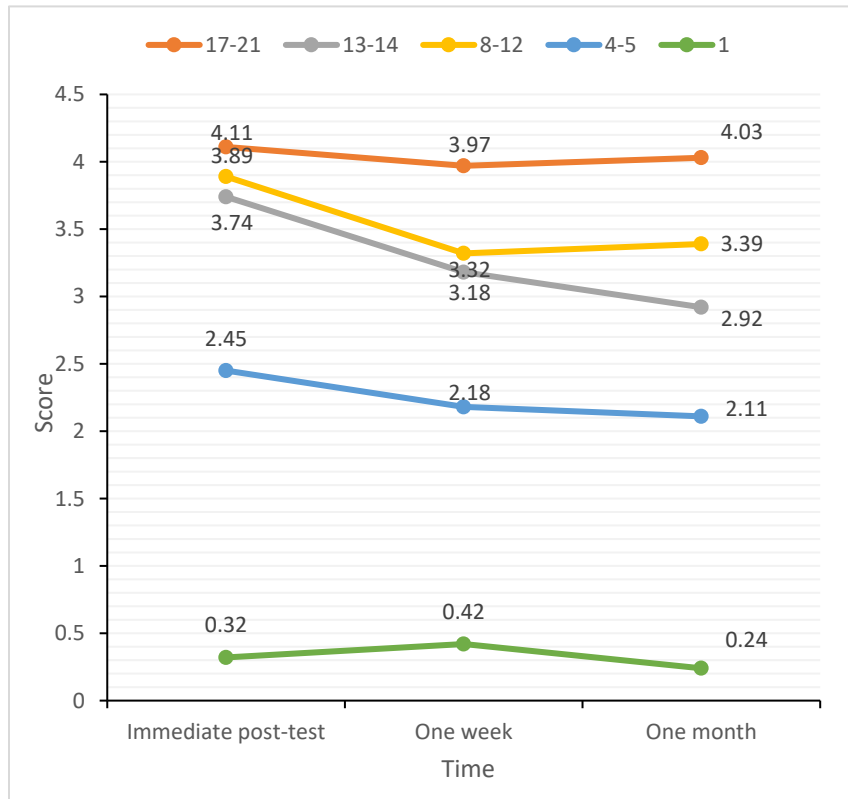
Words appearing more frequently seem to have a higher acquisition rate. Specifically, 4.11 out of 5 words that met 17-21 times were learnt over the treatment. For words that appeared 13-14 times and 8-12 times, the mean scores were 3.74 and 3.89, respectively. Words encountered 4-5 times had a mean score of 2.45. However, of the five words which occurred once, only 0.32 items were recognized. There was an overall decline in scores by frequency of exposure. Concerning the immediate post-test, there was a decrease in scores from 4.11 of the 5 test items (82.2%) for the 17-21 group to 0.32 (6.4%) for words which appeared once. Similar patterns were found after a week (3.97 to 0.42) and after a month (4.03 to 0.24). Words which were met 17-21 times had a mean score of around 4 for all three test administrations while items which occurred only once registered a score of near 0 on all those tests. From the result, it is likely that the less frequent items had lower (meaning) recognition rates than the more frequent items.

A Kruskal-Wallis H test indicated that frequency of exposure significantly influenced the mean score on the MC recognition test at the immediate post-test, $H(4) = 95.217$, $p = .000$. There were statistically significant differences in gain scores between encountering words that occurred 1 time and 4-5 times ($p = .000$, $r = -0.523$), 8-12 times ($p = .000$, $r = -0.895$), 13-14 times ($p = .000$, $r = -0.817$), and 17-21 times ($p = .000$, $r = -0.967$). Statistically significant differences in learning gains were also found between seeing words 4-5 times and 8-12 times ($p = .012$, $r = -0.372$), and 4-5 times and 17-21 times ($p = .001$, $r = -0.444$). Pairwise comparisons also showed that there were no significant differences in the mean scores of participants encountering items that occurred 13-14 times and 8-12 times ($p = 1.000$, $r = 0.078$), or 17-21 times ($p = 1.000$, $r = -0.150$). Moreover, there was no significant difference between the 8-12 group and 17-21 group ($p = 1.000$, $r = -0.072$).

Overall, results from Kruskal-Wallis test on all three test formats showed that one exposure acted differently from all other frequency bands with mean scores being comparatively low, at near zero. Meanwhile the 8-12, 13-14, and 17-21 frequency band acted in quite a similar way with learning gains being much higher. This seems to suggest that learning occurred when participants encountered a word a sufficient number of times (in this case: eight meetings), and it is not necessarily that the more repetitions, the higher the learning gains. This aligns with Webb and Nation (2017) claiming that there are words that are readily acquired after a very small number of repetitions, and some that are not, even after many encounters.

1.3 RQ3: Is there a relationship between participants' vocabulary learning proficiency indicated from the Guessing from Context Test and their vocabulary learning gains in the immediate test after reading a graded reader? If so, to what extent?

To answer research question 3, participants' scores on the GCT and the immediate post-tests were collected; these data were used to run a correlation analysis in SPSS.

TABLE 6

The mean scores and standard deviations (SD) of the Guessing from Context Test and the immediate post-tests

	Test	*Mean*	*SD*
GCT	Part 1 (Max= 20)	19.37	(1.00)
	Part 2 (Max= 20)	16.05	(2.28)
	Part 3 (Max= 20)	16.05	(1.74)
	Total (Max= 60)	51.47	(3.80)
Immediate post-tests (Max = 25)	Form Recognition	17.61	(2.80)
	Meaning Translation	8.67	(4.79)
	Multiple-choice Recognition	14.50	(4.03)

Note. n= 38.

TABLE 6 shows the mean scores of the GCT and the immediate post-tests. As shown in Table 6, the mean score for section 1 of the GCT is 19.37 (SD = 1.00). Since the maximum score for each part of the GCT is 20, there seems to be a ceiling effect in section 1 of this test. As a result, only the data from part 2, part 3 and the total score of the GCT were utilized to run the correlation analysis.

TABLE 7

Correlation analysis between GCT scores and learning gains in the immediate post-tests

		Form Recognition	Meaning Translation	Meaning Recognition
GCT2	Pearson Correlation	-.043	.137	.047
	Sig.	.797	.413	.779
GCT3	Pearson Correlation	-.090	-.111	-.150
	Sig.	.592	.505	.368
GCT Total	Pearson Correlation	-.114	.026	-.062
	Sig.	.496	.877	.712

Note. n= 38.

*. Correlation is significant at the 0.05 level (2-tailed)

A Pearson Correlation coefficient was computed to assess the relationship between participants' vocabulary learning proficiency indicated from the GCT and their vocabulary learning gains on the immediate post-tests (See TABLE 7). From the results there was no statistically significant correlation between these two variables across all parts of the GCT and the whole test. To be specific, there was a negative correlation between the score of GCT part 2 and that of the immediate word-form recognition test, $r = -.043$, $p = .797$. As for the immediate meaning translation test and the MC recognition test, positive correlations were found. The former produced $r = .137$, $p = .413$, whereas the latter registered $r = .047$, $p = .779$. However, these relationships were found not to be statistically significant. Concerning part 3 of the GCT, there were negative correlations between the score of GCT and all test types of the immediate post-test: form recognition, $r = -.090$, $p = .592$; meaning translation, $r = -.111$, $p = .505$; meaning recognition, $r = -.150$, $p = .368$. Overall, although there were some positive and some negative relationships, no statistically significant correlations were found, indicating that guessing meaning from context is not related to the extent of learning that occurs from all three tests (form recognition, meaning recall and meaning recognition).

2. Discussion

2.1 Learning gains concerning different test formats

Overall, participants were able to learn new words from reading a graded reader, with each test type yielding different gain scores. At the immediate post-test, participants' scores on the word-form recognition test were the highest (17.61/25 words) (70.44%), followed by the MC recognition test (14.5/25) (58%) and then the meaning-translation test (8.67/25) (34.68%). The results also showed that comparatively little decay occurred over one month in terms of the recognition tests, with reported decay rate being at around 12%. Concerning the meaning-translation test, about 40% of the learning gains were lost after one month. The extent of the learning gains varies with the specific measures used. It can be seen that the form recognition test resulted in the highest gain scores and meaning recall test produced the lowest. This is broadly in line with previous research which has shown that recognition tests are easier than recall instruments (González-Fernández and Schmitt 2019; Laufer and Goldstein 2004), and mirrors the findings of Waring and Takaki (2003).

It is important to remember that vocabulary learning is an incremental process and these different measures tap into differing strengths of knowledge. According to Barclay and Schmitt (2019), word learning is incremental, in which several aspects of word knowledge are mastered at different rates and the knowledge of different components of a word will be gained to varying degrees. Moreover, as preceding studies have pointed out, it is essential that several tests at different sensitivities are used to measure gains in lexical knowledge (Nation and Webb 2011; Pigada and Schmitt 2006; Webb 2007, 2008). It is explained that were only one comparatively challenging measure used (e.g., a translation test), then the amount of learning taking place would be greatly underestimated since it did not represent the strength of knowledge typically obtained from incidental learning. The fact that the form recognition and meaning recognition tests produced greater gain scores and retention rates might suggest that some partial knowledge not accessed by the meaning-translation test was found to be known via tests where participants' knowledge was prompted. Having three different tests like this could, therefore, demonstrate a range of vocabulary learning that took place from reading. Thus, it is still essential to include different test formats when conducting research of this nature.

2.2 Number of encounters needed to learn a word

Studies to date have attempted to determine the frequency of occurrence needed for learning to occur. The results of this study suggest that the number of encounters needed for considerable learning to take place is likely to be at least eight times. Meaning-translation test results showed that having met a word 8-12 times can result in about a 52% chance of that word being learned. The chance of learning a word rises to 60% when the word is encountered 17-21 times. However, there is only a 20-35% chance that a word's meaning will be remembered after one month, even if it was encountered more than 21 times. If the word was met only once, the chance is near zero. The results support Waring and Takaki's (2003) findings and other studies (Brown, Waring and Donkaewbua 2008; Pellicer-Sánchez and Schmitt 2010; Webb 2007) that (1) participants are more successful in learning words from context if they encounter the words several times: the more frequently an item is encountered, the more chance it has of being acquired and (2) unless words are reencountered soon after reading and with a sufficient number of times, then decay of word knowledge gained is likely to occur.

As indicated from the data, the vocabulary retention rate seems to be low. One reason for this might be that participants had too few chances to learn the words. Even items encountered more than 21 times in the text still have only a 60% chance of being acquired, let alone one meeting. It seems to suggest that in order for most of those words to be learnt, it would take much higher than 21 meetings. In alignment with this, it is believed that a sufficient number might be nearly 30-50 times or higher for new words encountered through graded reading (Waring 2008, see Brown, Waring and Donkaewbua 2008). Should learners read at 96-99% coverage, and it takes more than 21 encounters with a word to acquire it as the data seem to illustrate, then learners might have to read several hundred, even thousand words in order to learn one new word from reading (ibid.). When compared to what is often recommended for L2 learners when reading graded readers (e.g., Nation and Wang (1999) suggested one book a week at the appropriate level), such amount of reading is considerable. Moreover, the higher the reading ability level is, the greater volume of text learners have to read to meet an unknown word (Nation 2013). It is because rarer words would be encountered less frequently; as a result, more text has to be read to encounter an unknown word the required number of times (ibid.). That said, vocabulary incidental learning is still a very essential aspect of learners' vocabulary training. In spite of small incidental learning gains, such gains are meaningful and crucial to learners' L2 lexical development because as noted earlier, vocabulary learning is an incremental process, and much can be learned about a word through repeated encounters. As Webb (2007) stated, through repeatedly meeting partially known or unknown words in input, knowledge of not only form-meaning connection, but also other aspects of lexical knowledge (e.g., grammatical functions, collocations) are likely to be gained.

Another reason might be due to the lack of conscious attention to the words. According to Schmidt's (1990) Noticing Hypothesis, input does not become intake for language learning if it is not noticed. Schmidt (2001) believed that more noticing might lead to more learning and those linguistic forms in input that are consciously attended to are more likely to be learnt than those that are not noticed. However, in the present study, participants are presumably focused on understanding and taking pleasure in the story rather than on the words themselves when reading the graded reader. As the nonwords were left unmarked for natural reading, participants

were not forced to notice these words; thus, their awareness of these test items might not be raised, potentially resulting in a low rate of vocabulary acquisition.

These reasons might suggest two things. First, for higher rate of acquisition to occur, learners should be encouraged to read a larger amount of text to increase repetition since reading one graded reader might have been insufficient. This aligns with Webb and Chang (2015b) proposing that the proportion of words acquired through reading was higher providing that learners read multiple texts (rather than a single text). As they explained, reading multiple texts positively impacts learners' lexical knowledge since it provides more opportunities for repetition and, in turn, strengthens knowledge of novel words. Second, a certain amount of consciousness should be paid to words to increase the likelihood of them being learnt. This can be done by highlighting words in texts, which is a form of textual enhancement. As Schmidt (2001) claimed, although noticing does not guarantee learning, it does make acquisition possible. In agreement with this, Barcroft (2009) and Webb and Boers (2013) believed that what we learn is largely determined by what we focus on. According to Webb and Nation (2017), highlighting words in input may result in some gains in knowledge of the word form through deliberate noticing. Bishop (2004) also showed that highlighting multi-word combinations increases the chance that they will be noticed and looked up in the dictionary. In fact, there have been many series of graded readers (e.g., Oxford University Press' Dominoes series) incorporating extensive reading with an intensive reading approach by highlighting words in the texts.

2.3 Guessing meaning from context and vocabulary acquisition

The data indicated that there is no statistically significant correlation between subjects' vocabulary learning proficiency indicated from the GCT and their learning gains in the immediate post-tests. This is a surprising result given that vocabulary scholars often believe that being good at guessing from context leads to better vocabulary acquisition (Sasao 2019; Webb and Nation 2017). The result can be interpreted in two ways.

First, there might exist a relationship between guessing from context and learning; however, the GCT might not be sensitive enough to measure differences in guessing from context (i.e., the test might be too easy), making it impossible to correlate that variable against the

learning gains. Accordingly, it is necessary that a better test be introduced to explain any variance in participants' scores, allowing the researcher to understand the relationship.

Another interpretation is that there is not any connection between guessing from context and learning. This might have significant consequences on the construct validity of the GCT because the GCT is designed to predict how well students can guess meaning from context. The fact that there is not any relationship might suggest that (1) guessing meaning from context does not seem to be related to the ability to learn words from context, and (2) guessing from context and learning from context might be two different constructs that need to be identified separately. This result appears to be quite significant because it goes against what researchers often believe; specifically, there seems to be an assumption among vocabulary studies that having higher competency in guessing meaning from context leads to more vocabulary learning (Webb and Nation 2017). Sasao (2019, p. 422), in a chapter discussing 'Measuring the ability to learn words', also claimed that 'vocabulary learning ability refers to the ability to increase vocabulary learning and may include at least six types of knowledge and skills'; one of which is guessing from context. The present study results, however, pointed out that it is not the case, and more research is needed to make these conclusions.

Chapter summary

This chapter addressed three RQs by discussing results based on quantitative analyses. There appears to be a great deal of similarity between the current study and the original of Waring and Takaki (2003). Results showed that (1) words can be learnt from reading a graded reader, but many of them were soon forgotten after one month, (2) words that are frequently met were more likely to be acquired, and (3) no statistically significant correlation was found between guessing from context and vocabulary learning gains. In the final chapter, the research findings and discussion will be summarized. Then, some limitations of the study and pedagogical and research implications will be introduced.

CHAPTER V: CONCLUSION

1. Summary of the findings and discussion

The results showed that subjects in this study were able to learn new words incidentally from reading the graded reader *A Little Princess*, with each test format producing different learning gains. Participants' scores reflected the tests' order of difficulty, with immediate post-test scores of 17.61 (word-form recognition), 14.50 (MC recognition), and 8.67 (meaning translation). This finding is supported by González-Fernández and Schmitt (2019) and Laufer and Goldstein (2004), who found that receptive knowledge of meaning and form is easier than productive knowledge of the same two lexical knowledge types. The data also suggested that there was generally robust retention on the form recognition and meaning recognition tests as only about 12% of the learning gains were lost over a month. On the contrary, of the 8.67 words that were acquired on the immediate translation test, only 5.2 words were retained over the one-month retention interval, which is a decay rate of around 40%. In terms of frequency of exposure, the findings mirror Waring and Takaki (2003) and other studies (Brown, Waring and Donkaewbua 2008; Pellicer-Sánchez and Schmitt 2010; Webb 2007) that words appearing more frequently in the text were more likely to be learnt than those appearing less frequently. The minimum number of encounters for considerable learning to occur seems to be eight times. As indicated from the meaning recall test, having met an item 8-12 times leads to a 52% chance of that item being acquired, and this rises to 60% when it is encountered 17-21 times. For words that appeared only once, there is little or no chance that these words would be learnt and retained. Thus, a massive amount of graded reading is needed to increase repetition, making it possible for new vocabulary to be learnt. Additionally, no statistically significant correlation was found between participants' ability to guess meaning from context and their vocabulary learning gains. This appears to be a surprising result as it contradicts what vocabulary scholars often believe (i.e., being good at guessing from context results in more vocabulary learning). However, it might also be the case that the GCT is not sensitive enough to measure differences in guessing from context, making it impossible to correlate that variable against learning gains. Therefore, further research is needed to confirm this observation, allowing researchers and readers to understand more about the relationship between guessing from context and vocabulary acquisition.

2. Limitations of the study

One of the limitations of this research is the small sample size. In this study, data from only 38 Vietnamese participants were analysed. It would be better to carry out the same study with more subjects and from other language backgrounds so as to create more reliable findings with a more accurate representation of data.

As this study is a replication of Waring and Takaki (2003), some limitations of the present research would also be limitations of the original one. A limitation identified in this research relates to the scoring scale of the meaning translation test. Specifically, learners could score 0, 0.5, or 1 depending on the precision of their answer, with 0.5 being awarded when participants provided a word that has a similar meaning to the correct answer. According to Nation and Web (2011), when a scale is utilized to measure lexical knowledge, the scoring procedure can give credit for partial knowledge. However, it seems to be difficult to understand the parameters of a score of 0.5, and there is room for subjectivity in awarding partial knowledge in this way. In agreement with this, Schmitt (2010) argued that such scoring scale involved some level of assessors' subjectivity (e.g., was participants' translation 'close enough' to show their knowledge of the word's meaning?). Therefore, the final score of the translation test might come from the assessor's judgment of the test-takers' response. Schmitt (2010, p. 221) further recommended that "it is always desirable to have a direct, consistent, and unambiguous relationship between learner output on a test and the scoring interpretation of that output". Thus, from an assessment perspective, it would be more reliable if there are just correct and incorrect answers (1 and 0).

An additional limitation is that the study did not control for word length concerning the use of nonwords. As Barclay and Pellicer-Sánchez (2021) explained, length can be operationalised as the number of letters, phonemes, and syllables that a word contains. Since 25 nonwords were taken from the study of Waring and Takaki (2003), no change had been made to the spelling of these test items, with words having different length (number of letters), ranging from four to eight letters. Research has shown that word length is one of the factors affecting learning burden, with shorter words found to be generally easier to learn (Laufer 1997; Schmitt 2010). In line with previous studies, Barclay and Pellicer-Sánchez (2021) found that length was associated with learning difficulty, with the longer words posing a greater burden than shorter ones. Since word

length was not controlled in this study, the vocabulary acquisition might have been affected because participants might find longer words (e.g., *brenches*) harder to learn than shorter items (e.g., *mork*), potentially impacting the final gain scores. Thus, it is recommended that researchers take this intralexical variable into account when conducting this kind of research in the future.

3. Pedagogical and research implications

Despite certain limitations, this study still has important pedagogical and research implications.

Results of this study showed that repetition is vital for lexical knowledge development, which is broadly in line with previous research suggesting that the more frequent a word is encountered, the stronger the knowledge of it will be (Pellicer-Sánchez and Schmitt 2010; Waring and Takaki 2003; Webb 2007). For this reason, L2 learners should be provided with sufficient opportunities to revise learnt lexical items. Moreover, as repetition is said to be chiefly related to quantity of input (Webb and Nation 2017), teachers should encourage students to read more by using series of graded readers (rather than just a single one), or incorporating activities such as the linked skills activity (i.e., the same content is dealt with across the four language skills) and repeated reading or viewing. However, it should be remembered that having a high number of exposures to a target item does not guarantee learning to occur. Therefore, other pedagogical strategies that increase learners' noticing of the target items might be useful. Some of them might be highlighting the target words in the text and looking up words in a dictionary or glossary (ibid.).

Additionally, this research has pointed out that there is no statistically significant correlation between vocabulary acquisition and the capacity of guessing meaning from context. This finding appears to contradict vocabulary scholars' belief (Sasao 2019; Webb and Nation 2017) assuming that learners' ability to guess from context has a positive influence on vocabulary acquisition. Therefore, a replication of my research would be welcomed so that we can have a better understanding of the relationship between these two variables. It cannot be denied that guessing from context is important. However, from this research finding, researchers, scholars and teachers might have to consider in what way it is crucial; whether it is an effective strategy for vocabulary learning or for reading comprehension. Moreover, a correlation analysis was conducted in this study; hence, further research should consider other statistical measures that

might be more robust (e.g., regression or multiple regression analysis) to investigate the relationship between guessing from context and vocabulary learning. Overall, however, this study has added to our understanding of the impact of frequency of exposure and guessing from context on lexical acquisition and identified potentially important avenues for future research.

REFERENCE LIST

Barclay, S. and Pellicer-Sánchez, A., 2021. Exploring the learning burden and decay of foreign language vocabulary knowledge: The effect of part of speech and word length. *ITL-International Journal of Applied Linguistics* [online]. DOI: https://doi.org/10.1075/itl.20011.bar [Accessed 20 August 2021].

Barclay, S. and Schmitt, N., 2019. Current perspectives on vocabulary teaching and learning. In: Gao, X., ed. *Second Handbook of English Language Teaching* [online]. Switzerland: Springer, 2019, pp. 799-819. Available at: https://link.springer.com/content/pdf/10.1007/978-3-030-02899-2_42.pdf [Accessed 25 September 2021].

Barcroft, J., 2009. Effects of synonym generation on incidental and intentional L2 vocabulary learning during reading. *TESOL Quarterly* [online]. 43 (1), 79–103. DOI: https://doi.org/10.1002/j.1545-7249.2009.tb00228.x [Accessed 28 September 2021].

Barcroft, J., 2015. *Lexical input processing and vocabulary learning* [eBook]. Amsterdam: John Benjamins. Available via: NTU library [Accessed 15 July 2021].

Beck, I.L., McKeown, M.G., and McCaslin, E.S., 1983. Vocabulary development: All contexts are not created equal. *The Elementary School Journal* [online]. 83 (3), 177-181. Available at: https://www.journals.uchicago.edu/doi/abs/10.1086/461307 [Accessed 9 September 2021].

Biber, D. and Conrad, S., 2001. Register variation: A corpus approach. In: Schiffrin, D., Tannen, D., and Hamilton, H.E., eds. *The handbook of discourse analysis* [eBook]. Malden, Mass.: Blackwell Publishers, 2001, pp. 175-196. Available via: NTU library [Accessed 14 July 2021].

Bishop, H., 2004. The effect of typographic salience on the look up and comprehension of unknown formulaic sequences. In: Schmitt, N., ed. *Formulaic sequences: acquisition, processing, and use* [eBook]*.* Amsterdam: John Benjamins, 2004, pp. 227–248. Available via: NTU library [Accessed 28 September 2021].

Brown, R., Waring, R. and Donkaewbua, S., 2008. Incidental vocabulary acquisition from reading, reading-while-listening, and listening to stories. *Reading in a Foreign Language* [online]. 20 (2), 136–163. Available at: http://hdl.handle.net/10125/66816 [Accessed 24 August 2021].

Burns, N. and Grove, S.K., 2001. *The practice of nursing research: Conduct, critique & utilization.* 4th ed. Philadelphia: Saunders

Chen, C. and Truscott, J., 2010. The Effects of Repetition and L1 Lexicalization on Incidental Vocabulary Acquisition. *Applied Linguistics* [online]. 31, (5), 693–713. DOI: https://doi.org/10.1093/applin/amq031 [Accessed 15 July 2021].

Clarke, D.F. and Nation, I.S.P., 1980. Guessing the meanings of words from context: Strategy and techniques. *System* [online]. 8 (3), 211-220. DOI: https://doi.org/10.1016/0346-251X(80)90003-2 [Accessed 24 August 2021].

Coady, J., 1997. L2 vocabulary acquisition: A synthesis of the research. In: Coady, J. and Huckin, T., eds. *Second language vocabulary acquisition: A rationale for pedagogy* [eBook]. Cambridge, UK: Cambridge University Press, 1997, pp. 273-290. Available via: NTU library [Accessed 1 August 2021].

Cobb, T., n.d. *Web Vocabprofile*. Available at: https://www.lextutor.ca/vp/ [Accessed 6 August 2021].

Cumming, A., 1990. Metalinguistic and Ideational Thinking in Second Language Composing. *Written Communication* [online]. 7 (4), 482–511. DOI: 10.1177/0741088390007004003 [Accessed 14 July 2021].

Daneman, M. and Green, I., 1986. Individual differences in comprehending and producing words in context*. Journal of Memory and Language* [online]*. 25 (1), 1–18. DOI: https://doi.org/10.1016/0749-596X(86)90018-5 [Accessed 14 June 2021].

Day, R. and Bamford, J., 2002. Top ten principles for teaching extensive reading. *Reading in a Foreign Language* [online]. 14 (2), 136–141. Available at: http://hdl.handle.net/10125/66761 [Accessed 1 August 2021].

Day, R., Omura, C., and Hiramatsu, M., 1991. Incidental EFL vocabulary learning and reading. *Reading in a Foreign Language* [online]. 7 (2), 541-551. Available at: https://scholarspace.manoa.hawaii.edu/bitstream/10125/67035/7_2_10125_67035_rfl72day.pdf [Accessed 15 August 2021].

Day, R.R. and Bamford, J., 1998. Extensive reading in the second language classroom. Cambridge, UK: Cambridge University Press.

de Groot, A.M.B. and Keijzer, R., 2000. What is hard to learn is easy to forget: The roles of word concreteness, cognate status, and word frequency in foreign-language vocabulary learning and forgetting. *Language Learning* [online]. 50 (1), 1-56. DOI: https://doi.org/10.1111/0023-8333.00110 [Accessed 20 August 2021].

Dörnyei, Z., 2007. *Research methods in applied linguistics: quantitative, qualitative, and mixed methodologies*. Oxford: Oxford University Press.

Dupuy, B. and Krashen, S., 1993. Incidental vocabulary acquisition in French as a foreign language. *Applied Language Learning* [online]. 4, 55-63. Available at: https://www.researchgate.net/profile/Beatrice_Dupuy/publication/234637696_Incidental_Vocabulary_Acquisition_in_French_as_a_Foreign_Language/links/574868c008ae5298d8818ee8/Incidental-Vocabulary-Acquisition-in-French-as-a-Foreign-Language.pdf [Accessed 15 August 2021].

Elgort, I., 2011. Deliberate learning and vocabulary acquisition in a second language. *Language Learning* [online]. 61 (2), 367-413. DOI: https://doi.org/10.1111/j.1467-9922.2010.00613.x [Accessed 14 July 2021].

Ellis, R. and Heimbach, R., 1997. Bugs and birds: children's acquisition of second language vocabulary through interaction. *System* [online]. 25 (2), 247-259. DOI: https://doi.org/10.1016/S0346-251X(97)00012-2 [Accessed 14 July 2021].

Ellis, R., 1994. Factors in the incidental acquisition of second language vocabulary from oral input: A review essay. *Applied Language Learning*. 5 (1), 1–32.

Ellis, R., 1999. *Learning a second language through interaction* [eBook]. Amsterdam: John Benjamins. Available via: NTU library [Accessed 15 July 2021].

Fahs, P.S., Morgan, L.L., and Kalman, M., 2003. A call for replication. *Journal of Nursing Scholarship* [online]. 35 (1), 67-72. DOI: https://doi.org/10.1111/j.1547-5069.2003.00067.x [Accessed 15 August 2021].

Field, A.P., 2013. *Discovering statistics using IBM SPSS* [eBook]. 4th ed. London: SAGE. Available via: NTU library [Accessed 26 August 2021].

Fraser, C.A., 1999. Lexical processing strategy use and vocabulary learning through reading. *Studies in Second Language Acquisition* [online]. 21 (2), 225–241. DOI: https://doi.org/10.1017/S0272263199002041 [Accessed 24 August 2021].

Godfroid, A., Ahn, J., Choi, I., Ballard, L., Cui, Y., Johnston, S., Lee, S., Sarkar, A., and Yoon, H.J., 2018. Incidental vocabulary learning in a natural reading context: An eye-tracking study. *Bilingualism: Language and Cognition* [online]. 21 (3), 563-584. DOI: https://doi.org/10.1017/S1366728917000219 [Accessed 24 August 2021].

González-Fernández, B. and Schmitt, N., 2019. Word Knowledge: Exploring the Relationships and Order of Acquisition of Vocabulary Knowledge Components. *Applied Linguistics* [online]. 41 (4), 481–505. DOI: https://doi.org/10.1093/applin/amy057 [Accessed 24 September 2021].

Haladyna, T.M. and Rodriguez, M.C., 2013. *Developing and Validating Test Items* [eBook]. New York: Routledge. Available via: NTU library [Accessed 1 August 2021].

Hall, S.J., 1991. Using split information tasks to learn mathematics vocabulary. *Guidelines* [online]. 14 (1), 72–77. Available at: https://www.researchgate.net/profile/Stephen_Hall29/publication/234690485_Using_Split_Inf

ormation_Tasks_to_Learn_Mathematics_Vocabulary/links/5a17695fa6fdcc50ade61074/Using-Split-Information-Tasks-to-Learn-Mathematics-Vocabulary.pdf [Accessed 25 July 2021].

Hayes, D.P., 1988. Speaking and writing: Distinct patterns of word choice. *Journal of Memory and Language* [online]. 27 (5), 572–585. DOI: https://doi.org/10.1016/0749-596X(88)90027-7 [Accessed 23 July 2021].

Henriksen, B., 1999. Three dimensions of vocabulary development. *Studies in Second Language Acquisition* [online]. 21 (2), 303-317. Available at: http://www.jstor.org/stable/44486441 [Accessed 25 September 2021].

Horst, M., Cobb, T., and Meara, P., 1998. Beyond a Clockwork Orange: Acquiring second language vocabulary through reading. *Reading in a Foreign Language* [online]. 11 (2), 207– 223. Available at: https://scholarspace.manoa.hawaii.edu/bitstream/10125/66953/11_2_10125_66953_rfl112horst-sm.pdf [Accessed 15 July 2021].

Hu, M. and Nation, I.S.P., 2000. Unknown Vocabulary Density and Reading Comprehension. *Reading in a Foreign Language* [online]. 13 (1), 403-430. Available at: https://scholarspace.manoa.hawaii.edu/bitstream/10125/66973/13_1_10125_66973_rfl131hsuehchao.pdf [Accessed 9 August 2021].

Huckin, T. and Coady, J., 1999. Incidental vocabulary acquisition in a second language: A review. *Studies in Second Language Acquisition* [online]. 21 (2), 181–193. DOI: https://doi.org/10.1017/S0272263199002028 [Accessed 1 August 2021].

Huckin, T., 1995. Critical discourse analysis. *The Journal of TESOL.* 2 (3), 95–112.

Hulstijn, J., 2001. Intentional and incidental second-language vocabulary learning: A reappraisal of elaboration, rehearsal and automaticity. In: Robinson, P., ed. *Cognition and second language instructions* [online]. Cambridge, UK: Cambridge University Press, 2001, pp. 258– 286. DOI: https://doi.org/10.1017/CBO9781139524780.011 [Accessed 15 July 2021].

Jenkins, J.R., Stein, M.L., and Wysocki, K., 1984. Learning vocabulary through reading. *American Educational Research Journal* [online]. 21 (4), 767–787. DOI: 10.3102/00028312021004767. [Accessed 15 July 2021].

Joe, A., 1998. What effects do text-based tasks promoting generation have on incidental vocabulary acquisition? *Applied Linguistics* [online]. 19 (3), 357–377. DOI: https://doi.org/10.1093/applin/19.3.357 [Accessed 25 July 2021].

Krashen, S., 1985. *The Input Hypothesis: Issues and Implications.* London: Longman.

Krashen, S., 1989. We acquire vocabulary and spelling by reading: additional evidence for the Input Hypothesis. *The Modern Language Journal* [online]. 73 (4), 440–464. DOI: 10.2307/326879 [Accessed 1 August 2021].

Kuppens, A.H., 2010. Incidental foreign language acquisition from media exposure. *Learning, Media and Technology* [online]. 35 (1), 65–85. DOI: https://doi.org/10.1080/17439880903561876 [Accessed 23 July 2021].

Laufer, B. and Goldstein, Z., 2004. Testing vocabulary knowledge: size, strength, and computer adaptiveness. *Language Learning* [online]*.* 54 (3)*,* 399–436. DOI: https://doi.org/10.1111/j.0023-8333.2004.00260.x [Accessed 14 July 2021].

Laufer, B. and Hadar, L., 1997. Assessing the effectiveness of monolingual, bilingual, and 'bilingualised' dictionaries in the comprehension and production of new words. *Modern Language Journal* [online]*.* 81 (2)*,* 189–196. DOI: https://doi.org/10.1111/j.1540-4781.1997.tb01174.x [Accessed 14 June 2021].

Laufer, B. and Ravenhorst-Kalovski, G.C., 2010. Lexical threshold revisited: Lexical text coverage, learners' vocabulary size and reading comprehension. *Reading in a Foreign Language* [online]. 22 (1), 15-30. DOI: 10125/66648 [Accessed 5 August 2021].

Laufer, B., 1989. What percentage of text-lexis is essential for comprehension? In: Lauren, C. and Nordman, M., eds. *Special language: From humans to thinking machines*. Clevedon: Multilingual Matters, 1989, pp. 316-323.

Laufer, B., 1997. What's in a word that makes it hard or easy? Intralexical factors affecting the difficult of vocabulary acquisition. In: Schmitt, N. and McCarthy, M., eds. *Vocabulary: Description, Acquisition, and Pedagogy.* Cambridge: Cambridge University Press, 1997, pp. 140-155.

Laufer, B., 2003. Vocabulary acquisition in a second language: do learners really acquire most vocabulary by reading? *Canadian Modern Language Review* [online]. 59 (4)*,* 565–585. DOI: 10.3138/cmlr.59.4.567 [Accessed 14 July 2021].

Laufer, B., 2005. Focus on form in second language vocabulary learning. In: Foster-Cohen, S.H., Garcia-Mayo, M., and Cenoz, J., eds. *EUROSLA yearbook* [online]. Amsterdam: Benjamins, 2005, pp. 223–250. DOI: https://doi.org/10.1075/eurosla.5.11lau [Accessed 25 July 2021].

Laufer, B., 2010. Form-focused instruction in second language vocabulary learning. In: Chacón-Beltrán, R., Abello-Contesse, C., and Torreblanca-López, M.M., eds. *Insights into non-native vocabulary teaching and learning* [eBook]. Bristol, Buffalo: Multilingual Matters, 2010, pp. 15-27. Available via: NTU library [Accessed 28 August 2021].

Leki, I. and Carson, J.G., 1994. Students' perceptions of EAP writing instruction and writing needs across the disciplines. *TESOL quarterly* [online]. 28 (1), 81-101. DOI: https://doi.org/10.2307/3587199 [Accessed 14 July 2021].

Lewis, M. 1993. *The Lexical Approach.* Hove, England: Language Teaching Publications.

Li, S., 2016. The Construct Validity of Language Aptitude: A Meta-Analysis. *Studies in Second Language Acquisition* [online]. 38 (4), 801-842. DOI: https://doi.org/10.1017/S027226311500042X [Accessed 20 August 2021].

Lindgren, E. and Muñoz, C., 2013. The influence of exposure, parents, and linguistic distance on young European learners' foreign language comprehension. *International Journal of*

Multilingualism [online]. 10 (1), 105–129. DOI: https://doi.org/10.1080/14790718.2012.679275 [Accessed 23 July 2021].

Lindstromberg, S., 2019. Intentional L2 vocabulary learning. In: Webb, S., ed. *The Routledge handbook of vocabulary studies* [eBook]. London: Routledge, 2019, pp. 240-254. Available via: NTU library [Accessed 15 July 2021].

Liu, N. and Nation, I.S.P., 1985. Factors affecting guessing vocabulary in context. *RELC Journal* [online]. 16 (1), 33–42. DOI: https://doi.org/10.1177/003368828501600103 [Accessed 25 July 2021].

Ludwig, J., 1984. Vocabulary acquisition as a function of word characteristics. *The Canadian Modern Language Review* [online]. 40 (4), 552-562. DOI: 10.3138/cmlr.40.4.552 [Accessed 14 July 2021].

Nagy, W.E. and Herman, P.A., 1987. Breadth and depth of vocabulary knowledge: implications for acquisition and instruction. In: McKeown, M.G. and Curtis, M.E., eds. *The nature of vocabulary acquisition*. Hillsdale, NJ: Erlbaum, 1987, pp. 19–36.

Nagy, W.E., Herman, P.A., and Anderson, R.C., 1985. Learning words from context. *Reading Research Quarterly* [online]. 20 (2), 233–253. DOI: 10.2307/747758. [Accessed 15 July 2021].

Nassaji, H., 2003. L2 vocabulary learning from context: Strategies, knowledge sources, and their relationship with success in L2 lexical inferencing. *TESOL Quarterly* [online]. 37 (4), 645–670. DOI: https://doi.org/10.2307/3588216 [Accessed 24 August 2021].

Nation, I.S.P. and Wang, K., 1999. Graded readers and vocabulary. *Reading in a Foreign Language* [online]. 12, 355-380. Available at: https://scholarspace.manoa.hawaii.edu/bitstream/10125/66966/12_2_10125_66966_rfl122nation.pdf [Accessed 19 August 2021].

Nation, I.S.P. and Webb, S., 2011. *Researching and analysing vocabulary*. Boston, M.A.: Heinle Cengage Learning.

Nation, I.S.P., 1990. *Teaching and learning vocabulary* [eBook]. Boston, Mass.: Heinle & Heinle. Available via: NTU library [Accessed 14 July 2021].

Nation, I.S.P., 1997. The language learning benefits of extensive reading. *The Language Teacher* [online]. 21 (5), 13–16. Available at: https://openaccess.wgtn.ac.nz/articles/journal_contribution/The_language_learning_benefits_of_extensive_reading/12560378/files/23412086.pdf [Accessed 1 August 2021].

Nation, I.S.P., 2006. How large a vocabulary is needed for reading and listening?. *Canadian modern language review* [online]. 63 (1), 59-82. DOI: https://doi.org/10.3138/cmlr.63.1.59 [Accessed 14 July 2021].

Nation, I.S.P., 2007. The four strands. *International Journal of Innovation in Language Learning and Teaching* [online]. 1 (1), 2-13. DOI: https://doi.org/10.2167/illt039.0 [Accessed 15 July 2021].

Nation, I.S.P., 2013. *Learning vocabulary in another language* [eBook]. 2nd edition. Cambridge: Cambridge University Press. Available via: NTU library [Accessed 20 July 2021].

Paribakht, T.S. and Wesche, M., 1997. Vocabulary enhancement activities and reading for meaning in second language vocabulary acquisition. In: Coady, J. and Huckin, T., eds. *Second language vocabulary acquisition: A rationale for pedagogy* [eBook]. Cambridge, UK: Cambridge University Press, 1997, pp. 174-200. Available via: NTU library [Accessed 25 July 2021].

Paribakht, T.S. and Wesche, M., 1999. Reading and "incidental" L2 vocabulary acquisition: An introspective study of lexical inferencing. *Studies in second language acquisition* [online]. 21 (2), 195-224. DOI: https://doi.org/10.1017/S027226319900203X [Accessed 5 August 2021].

Parry, K., 1991. Building a vocabulary through academic reading. *TESOL Quarterly* [online]. 25 (4), 629– 653. DOI: https://doi.org/10.2307/3587080 [Accessed 24 August 2021].

Pavia, N., Webb, S., and Faez, F., 2019. Incidental vocabulary learning through listening to songs. *Studies in Second Language Acquisition* [online]. 41 (4), 745-768. DOI: https://doi.org/10.1017/S0272263119000020 [Accessed 20 July 2021].

Pellicer-Sánchez, A. and Schmitt, N., 2010. Incidental vocabulary acquisition from an authentic novel: Do Things Fall Apart?. *Reading in a Foreign Language* [online]. 22 (1), 31-55. Available at: https://scholarspace.manoa.hawaii.edu/bitstream/10125/66652/1/22_1_10125_66652_pellicer sanchez.pdf [Accessed 28 July 2021].

Pellicer-Sánchez, A., 2016. Incidental L2 vocabulary acquisition from and while reading: an eye-tracking study. *Studies in Second Language Acquisition* [online]. 38 (1), 97–130. DOI: https://doi.org/10.1017/S0272263115000224 [Accessed 23 July 2021].

Pellicer-Sánchez, A., 2017. Learning L2 collocations incidentally from reading. *Language Teaching Research* [online]. 21 (3), 381–402. DOI: https://doi.org/10.1177/1362168815618428 [Accessed 15 July 2021].

Peters, E. and Webb, S., 2018. Incidental vocabulary acquisition through watching a single episode of L2 television. *Studies in Second Language Acquisition* [online]. 40 (3), 551–557. DOI: https://doi.org/10.1017/ S0272263117000407 [Accessed 20 July 2021].

Peters, E., 2018. The effect of out-of-class exposure to English language media on learners' vocabulary knowledge. *ITL International Journal of Applied Linguistics* [online]. 169 (1), 142–168. DOI: https://doi.org/10.1075/itl.00010.pet [Accessed 23 July 2021].

Peters, E., 2019. Factors Affecting the Learning of Single-Word Items. In: Webb, S., ed. *The Routledge Handbook of Vocabulary Studies* [eBook]. London: Routledge, 2019, pp. 125-142. Available via: NTU library [Accessed 20 August 2021].

Pham, H.H., 2005. University English classrooms in Vietnam. *ELT Journal* [online]. 59 (4), 336-338. DOI: https://doi.org/10.1093/elt/cci063 [Accessed 28 September 2021].

Pigada, M. and Schmitt, N., 2006. Vocabulary acquisition from extensive reading: A case study. *Reading in a Foreign Language* [online]. 18 (1), 1–28. Available at: https://eric.ed.gov/?id=EJ759833 [Accessed 1 August 2021].

Pitts, M., White, H., and Krashen, S., 1989. Acquiring second language vocabulary through reading: A replication of the Clockwork Orange study using second language acquirers. *Reading in a Foreign Language* [online]. 5 (2), 271–275. Available at: https://scholarspace.manoa.hawaii.edu/bitstream/10125/67017/5_2_10125_67017_rfl52pitts.pdf [Accessed 15 July 2021].

Polit, D.F. and Hungler, B.P., 1999. *Nursing research: Principles and Methods.* 6th ed. Philadelphia: Lippincott.

Porte, G., ed., 2012. *Replication Research in Applied Linguistics*. Cambridge: Cambridge University Press.

Read, J. and Nation, I.S.P., 1986. *Some Issues in the Testing of Vocabulary Knowledge.* [Paper presented at a language testing symposium, Israel]. 11-13 May. Available at: https://files.eric.ed.gov/fulltext/ED277282.pdf [Accessed 14 June 2021].

Read, J., 2000. *Assessing Vocabulary* [eBook]. Cambridge: Cambridge University Press. Available via: NTU library [Accessed 14 June 2021].

Richards, J., 1976. The Role of Vocabulary Teaching. *TESOL Quarterly* [online]. 10 (1), 77-89. DOI: 10.2307/3585941 [Accessed 25 September 2021].

Rivers, W.M. and Nunan, D., 1991. *Language teaching methodology: A textbook for teachers.* Hertfordshire: Prentice Hall International.

Rodgers, M.P. and Webb, S., 2019. Incidental vocabulary learning through viewing television. *ITL-International Journal of Applied Linguistics* [online]. 171 (2), 191-220. DOI: https://doi.org/10.1075/itl.18034.rod [Accessed 23 July 2021].

Rodgers, M.P.H. and Webb, S., 2011. Narrow viewing: the vocabulary in related television programs. *TESOL Quarterly* [online]. 45 (4), 689–717. DOI: https://doi.org/10.5054/tq.2011.268062 [Accessed 23 July 2021].

Rott, S., 1999. The effect of exposure frequency on intermediate language learners' incidental vocabulary acquisition through reading. *Studies in Second Language Acquisition* [online]. 21 (1), 589–619. DOI: https://doi.org/10.1017/S0272263199004039 [Accessed 24 August 2021].

Santos, T., 1989. Replication in Applied Linguistics Research. *TESOL Quarterly* [online]. 23 (4), 699-702. DOI: 10.2307/3587548 [Accessed 28 August 2021].

Saragi, T., Nation, I.S.P., and Meister, G.F., 1978. Vocabulary learning and reading. *System*. 6, 72–78.

Sasao, Y. and Webb, S., 2017. The Word Part Levels Test. *Language Teaching Research* [online]. 21 (1)*, 12–30. DOI: https://doi.org/10.1177/1362168815586083 [Accessed 14 June 2021].

Sasao, Y. and Webb, S., 2018. The guessing from context test. *ITL- International Journal of Applied Linguistics* [online]. 169 (1), 115-141. Available at: https://www.lextutor.ca/tests/sasao_webb_guessing_2018.pdf [Accessed 24 August 2021].

Sasao, Y., 2013. *Diagnostic tests of English vocabulary learning proficiency: guessing from context and knowledge of word parts.* Ph.D. thesis, Victoria University of Wellington, New Zealand.

Sasao, Y., 2019. Measuring the Ability to Learn Words. In: Webb, S., ed. *The Routledge handbook of vocabulary studies* [eBook]. London: Routledge, 2019, pp. 419-432. Available via: NTU library [Accessed 6 September 2021].

Schmidt, R.W., 1990. The role of consciousness in second language learning. *Applied linguistics* [online]. 11 (2), 129-158. DOI: https://doi.org/10.1093/applin/11.2.129 [Accessed 6 September 2021].

Schmidt, R.W., 2001. Attention. In: Robinson, P., ed. *Cognition and second language instruction*. New York: Cambridge University Press, 2001, pp. 3–32.

Schmitt, N. and McCarthy, M., 1997. *Vocabulary: Description, Acquisition and Pedagogy.* Cambridge: Cambridge University Press.

Schmitt, N. and Zimmerman, C.B., 2002. Derivative word forms: What do learners know? *TESOL Quarterly* [online]. 36 (2), 145–171. DOI: https://doi.org/10.2307/3588328 [Accessed 25 July 2021].

Schmitt, N., 1997. Vocabulary learning strategies. In: Schmitt, N. and McCarthy, M., eds. *Vocabulary: Description, Acquisition, and Pedagogy.* Cambridge: Cambridge University Press, 1997, pp. 199-227.

Schmitt, N., 2000. *Vocabulary in language teaching* [eBook]. Cambridge: Cambridge University Press. Available via: NTU library [Accessed 28 August 2021].

Schmitt, N., 2008. Instructed second language vocabulary learning. *Language Teaching Research* [online]. 12 (3), 329–363. DOI: 10.1177/1362168808089921 [Accessed 25 July 2021].

Schmitt, N., 2010. *Researching Vocabulary: A Vocabulary Research Manual* [eBook]. Basingstoke: Palgrave Macmillan. Available via: NTU library [Accessed 6 August 2021].

Schmitt, N., Schmitt, D., and Clapham, C., 2001. Developing and exploring the behaviour of two new versions of the Vocabulary Levels Test. *Language Testing* [online]. 18 (1), 55-88. DOI: https://doi.org/10.1177/026553220101800103 [Accessed 5 August 2021].

Schouten-van Parreren, C., 1996. Vocabulary learning and metacognition. In: Sajavaara, K. and Fairweather, C., eds. *Approaches to second language acquisition*. Jyvaskyla: University of Jyvaskyla, 1996, pp. 63–69.

Shanahan, T., 2017. Reading research: The importance of replication. *The Reading Teacher* [online]. 70 (4), 507-510. DOI: https://doi.org/10.1002/trtr.1520 [Accessed 15 August 2021].

Taris, T., 2000. *A primer in longitudinal data analysis* [eBook]. London: SAGE. Available via: NTU library [Accessed 5 August 2021].

Tinkham, T., 1993. The Effect of Semantic Clustering on the Learning of Second Language Vocabulary. *System* [online]. 21 (3), 371-380. DOI: https://doi.org/10.1016/0346-251X(93)90027-E [Accessed 20 August 2021].

van Zeeland, H. and Schmitt, N., 2013. Lexical coverage in L1 and L2 listening comprehension: the same or different from reading comprehension? *Applied Linguistics* [online]. 34 (4)*,* 457–479. DOI: https://doi.org/10.1093/applin/ams074 [Accessed 14 June 2021].

Vidal, K., 2011. A comparison of the effects of reading and listening on incidental vocabulary acquisition. *Language Learning* [online]. 61 (1), 219–258. DOI: https://doi.org/10.1111/j.1467-9922.2010.00593.x [Accessed 23 July 2021].

Waring, R. and Takaki, M., 2003. At what rate do learners learn and retain new vocabulary from reading a graded reader? *Reading in a Foreign Language* [online]. 15 (2), 130-163. Available at: https://scholarspace.manoa.hawaii.edu/bitstream/10125/66776/1/15_2_10125_66776_waring .pdf [Accessed 15 July 2021].

Webb, S. and Boers, F., 2013. *Do textual enhancement techniques increase incidental learning of collocation?* [Paper presented at Vocab@Vic conference, Victoria University of Wellington, New Zealand]. December.

Webb, S. and Chang, A.C.S., 2012. Vocabulary learning through assisted and unassisted repeated reading. *Canadian Modern Language Review* [online]. 68 (3), 1–24. DOI: https://doi.org/10.3138/cmlr.1204.1 [Accessed 23 July 2021].

Webb, S. and Chang, A.C.S., 2015a. How does prior word knowledge affect vocabulary learning progress in an extensive reading program? *Studies in Second Language Acquisition* [online]. 37 (4), 651–675. DOI: https://doi.org/10.1017/S0272263114000606 [Accessed 15 July 2021].

Webb, S. and Chang, A.C.S., 2015b. Second language vocabulary learning through extensive reading: How does frequency and distribution of occurrence affect learning? *Language Teaching Research* [online]. 19 (6), 667–686. DOI: 10.1177/1362168814559800 [Accessed 15 July 2021].

Webb, S. and Rodgers, M.P.H., 2009. The lexical coverage of movies. *Applied Linguistics* [online]. 30 (3), 407–427. DOI: https://doi.org/10.1093/applin/amp010 [Accessed 23 July 2021].

Webb, S., 2005. Receptive and productive vocabulary learning: The effects of reading and writing on word knowledge. *Studies in second language acquisition* [online]. 27 (1), 33-52. DOI: https://doi.org/10.1017/S0272263105050023 [Accessed 1 August 2021].

Webb, S., 2007. The effects of repetition on vocabulary knowledge. *Applied Linguistics* [online]. 28 (1), 46–65. DOI: https://doi.org/10.1093/applin/aml048 [Accessed 15 July 2021].

Webb, S., 2008. The effects of context on incidental vocabulary learning. *Reading in a Foreign Language* [online]. 20 (2), 232–245. Available at: https://files.eric.ed.gov/fulltext/EJ815123.pdf [Accessed 9 September 2021].

Webb, S., 2009. The effects of pre-learning vocabulary on reading comprehension and writing. *Canadian Modern Language Review* [online]. 65 (3), 441–470. DOI: 10.3138/cmlr.65.3.441 [Accessed 14 July 2021].

Webb, S., 2019. *The Routledge Handbook of Vocabulary Studies* [eBook]. London: Routledge. Available via: NTU library [Accessed 20 July 2021].

Webb, S., Newton, J., and Chang, A.C.S., 2013. Incidental learning of collocation. *Language Learning* [online]. 63 (1), 91–120. DOI: https://doi.org/10.1111/j.1467-9922.2012.00729.x [Accessed 15 July 2021].

Webb, S.A. and Nation, I.S.P., 2017. *How vocabulary is learned* [eBook]. Oxford: Oxford University Press. Available via: NTU library [Accessed 14 July 2021].

Wilkins, D.A., 1972. *Linguistics in language teaching*. London: Edward Arnold.

Zimmerman, C.B., 1997. Do reading and interactive vocabulary instruction make a difference? An empirical study. *TESOL Quarterly* [online]. 31 (1), 121–140. DOI: https://doi.org/10.2307/3587978 [Accessed 25 July 2021].

APPENDICES

Appendix 1. Vocabulary Levels Test (VLT) version 2 at 2000-word level (Schmitt, Schmitt and Clapham 2001)

1			2		
	1. copy			1. accident	
	2. event	_____ end or highest point		2. debt	_____ loud deep sound
	3. motor	_____ this moves a car		3. fortune	_____ something you
	4. pity	_____ thing made to be like		4. pride	must pay
	5. profit	another		5. roar	____ having a high
	6. tip			6. thread	opinion of yourself

3			4		
	1. coffee			1. clerk	
	2. disease	_____ money for work		2. frame	_____ a drink
	3. justice	_____ a piece of clothing		3. noise	_____ office worker
	4. skirt	_____ using the law in the		4. respect	_____ unwanted sound
	5. stage	right way		5. theatre	
	6. wage			6. wine	

5			6		
	1. dozen			1. admire	
	2. empire	_____ chance		2. complain	_____ make wider or
	3. gift	_____ twelve		3. fix	longer
	4. opportunity	_____ money paid to the		4. hire	_____ bring in for the first
	5. relief	government		5. introduce	time
	6. tax			6. stretch	_____ have a high opinion
					of someone

7			8		
	1. arrange			1. blame	
	2. develop	_____ grow		2. elect	_____ make
	3. lean	_____ put in order		3. jump	_____ choose by voting
	4. owe	_____ like more than		4. manufacture	_____ become like water
	5. prefer	something else		5. melt	
	6. seize			6. threaten	

9			10		
	1. ancient			1. bitter	
	2. curious	_____ not easy		2. independent	_____ beautiful
	3. difficult	_____ very old		3. lovely	_____ small
	4. entire	_____ related to God		4. merry	_____ liked by many
	5. holy			5. popular	people
	6. social			6. slight	

Appendix 2. Guessing from Context Test (Sasao 2013)

Available at:

https://ysasaojp.info/wp-content/uploads/2020/08/Vocabulary_Tests_GCT_GCT.pdf

Appendix 3. Post-tests

A. Word-form recognition test

Select the words that you met in the story *A Little Princess*

(Hãy chọn các từ mà bạn đã gặp khi đọc câu chuyện)

bandle	fale	bettle	bick	tantic	prink
bing	flart	windle	loncher	sind	mand
borch	mave	tance	tring	vack	toker
clath	nutious	parrow	cadle	chorm	palk
crasty	quent	greal	smort	blund	stoll
doce	sheddle	mear	molden	mork	rimple
diggle	smick	brenches	nase	yelt	speat

B. Meaning (translation) test

What do these words mean? Write the meaning in Vietnamese[2]. (Những từ dưới đây có nghĩa gì? Hãy viết nghĩa của chúng bằng tiếng Việt. Nếu bạn có nhiều hơn 1 ý tưởng, hãy viết theo thứ tự mà bạn tự tin.)

1	windle	1………………………………………	2………………………………………
2	loncher	1………………………………………	2………………………………………
3	mand	1………………………………………	2………………………………………
4	brenches	1………………………………………	2………………………………………
5	mear	1………………………………………	2………………………………………
6	mork	1………………………………………	2………………………………………
7	cadle	1………………………………………	2………………………………………
8	smort	1………………………………………	2………………………………………
9	tantic	1………………………………………	2………………………………………
10	bettle	1………………………………………	2………………………………………
11	parrow	1………………………………………	2………………………………………
12	chorm	1………………………………………	2………………………………………
13	molden	1………………………………………	2………………………………………
14	tring	1………………………………………	2………………………………………
15	toker	1………………………………………	2………………………………………
16	nase	1………………………………………	2………………………………………
17	bick	1………………………………………	2………………………………………

[2] To encourage responses, subjects were given two chances to answer.

18	prink	1…………………………………………	2…………………………………………
19	sind	1…………………………………………	2…………………………………………
20	greal	1…………………………………………	2…………………………………………
21	blund	1…………………………………………	2…………………………………………
22	palk	1…………………………………………	2…………………………………………
23	tance	1…………………………………………	2…………………………………………
24	vack	1…………………………………………	2…………………………………………
25	rimple	1…………………………………………	2…………………………………………

C. Multiple-choice recognition test

Choose the word with the nearest meaning. When you do not know an item, please choose 'I do not know'. (Chọn từ có nghĩa gần nhất với từ in đậm. Khi bạn không rõ về một từ nào đó, hãy chọn 'I do not know')

1	**blund**	sun	mountain	photo	flower	I do not know
2	**palk**	happy	doubtful	special	easy	I do not know
3	**tance**	air	moment	love	respect	I do not know
4	**vack**	hard	busy	free	wrong	I do not know
5	**rimple**	world	mouth	music	club	I do not know
6	**parrow**	letter	piano	hand	name	I do not know
7	**loncher/s**	teacher/s	nurse/s	servant/s	farmer/s	I do not know
8	**molden**	peaceful	hot	clean	dead	I do not know
9	**tring**	rich	dark	pretty	interesting	I do not know
10	**toker**	shoe	bread	car	stair	I do not know
11	**mork**	red	clever	mad	good	I do not know
12	**cadle**	tree	night	college	glass	I do not know

13	**smort**	dry	crazy	beautiful	dirty	I do not know
14	**tantic**	new	intelligent	cold	active	I do not know
15	**bettle**	cow	window	mud	station	I do not know
16	**nase**	bag	head	paper	desk	I do not know
17	**bick**	late	ugly	wet	exact	I do not know
18	**prink/s**	box/es	bike/s	week/s	hat/s	I do not know
19	**sind**	snow	pepper	chair	eye	I do not know
20	**greal**	paper	tape	game	winter	I do not know
21	**windle/s**	cup/s	elephant/s	house/s	book/s	I do not know
22	**chorm/s**	child/children	year/s	sea/s	bird/s	I do not know
23	**mand/s**	dog/s	room/s	face/s	sky/skies	I do not know
24	**brenches**	rings	mines	songs	cakes	I do not know
25	**mear**	money	pen	cat	file	I do not know

Appendix 4. List of L2 words and nonwords equivalent and the number of recurrences in the text

Number	L2 words	Nonwords	Number of recurrences in the text	Test word group
1	servant/s	loncher/s	20	17-21 Group
2	face/s	mand/s	18	
3	mines	brenches	18	
4	house/s	windle/s	17	
5	child/children	chorm/s	21	
6	money	mear	14	13-14 Group
7	good	mork	14	
8	night	cadle	13	
9	beautiful/ly	smort/ly	14	
10	window	bettle	14	
11	new	tantic	12	8-12 Group
12	name	parrow	9	
13	dead	molden	8	
14	rich/richer/richest	tring/tringer/tringest	11	
15	bread	toker	8	
16	head	nase	4	4-5 Group
17	late	bick	4	
18	week/s	prink/s	5	
19	snow	sind	4	
20	winter	greal	4	
21	sun	blund	1	One occurrence Group
22	special	palk	1	
23	moment	tance	1	
24	wrong	vack	1	
25	world	rimple	1	
			Total: 237 recurrences	

Adapted from Waring and Takaki (2003)

Appendix 5. Letter of Consent

Doan Trang Nguyen
MA in English Language Teaching
Nottingham Trent University

Dear Participant

I would like to invite you to participate in a study I am conducting as part of my MA degree in English Language Teaching at Nottingham Trent University under the supervision of Sam Barclay.

The title of my study is 'The relationship between guessing from context, frequency of exposure, and vocabulary acquisition: An investigation into Vietnamese English learners' incidental learning from reading a graded reader.'

The aim of my research is to examine the rate at which vocabulary is learned from reading a graded reader and retained over time and investigate the relationship between learners' ability to guess from context and learning gains.

First, you are asked to do the Vocabulary Levels Test (VLT) at 2,000-word level to see whether you have knowledge of the 2000 most frequent words. If achieving a score lower than 26/30, you are not qualified to continue the project. If the test score is 26/30 or higher, you will continue to do the Guessing from Context Test. After that, there will be a reading session in which you are asked to read a graded reader then finish a test immediately after reading. Due to the contextual constraint, the graded reader will be an electronic version. You will be tested again after a week, and after a month.

My data will only be used for academic research purposes. Your identity will be protected and your name will not appear in the study.

Would you please indicate your agreement to participate by signing the letter of consent below and returning it to me by email? If you have any questions regarding this study, or would like additional information to assist you in reaching a decision about participation, please do not hesitate to contact me.

Thank you.

Doan Trang Nguyen

Nottingham Trent University

Nottingham, NG1 5LP

N0952412@my.ntu.ac.uk

Letter of Consent

I agree to take part in the research study described above. I understand that the research will be presented as part of an MA dissertation.

I understand that any information I provide may be used in the study, but that my name will not appear and that every effort will be made to protect my identity.

I understand that my participation is voluntary, and that I can choose not to participate or to withdraw at any time. If I decide to withdraw, I agree to inform the researcher personally.

Name of Participant:

Signature of Participant:

Date: