

**The effects of L1 and L2 glosses in modified  
texts on incidental second language  
vocabulary acquisition**

**Dongyang Yan**

**MATESOL**

**University of York**

**Department of Education**

**September 2015**

## Abstract

The present study, a partial replication of the studies from Hulstijn (1992) and Jacobs et al. (1994), investigates the effectiveness of first language (L1) and second language (L2) glosses on incidental L2 vocabulary acquisition, as a by-product of reading. 24 Chinese college-level English bilinguals were assigned into three different groups including no gloss, L1 gloss, and L2 gloss groups. They were tested through a reading comprehension test as well as an immediate multiple-choice test and a delayed multiple-choice vocabulary test one week later. A questionnaire reporting their preferences and attitudes towards the use of glossing in reading materials was conducted after the immediate vocabulary test. One-way analysis of variance (ANOVA) and mix design repeated measure ANOVA were utilised to analyse the data obtaining from tests. The results from this study confirmed the effectiveness of glosses on reading comprehension and on vocabulary retention, suggesting that reading texts either with L1 or L2 glosses are beneficial for L2 learners. However, no significant difference was found between L1 and L2 gloss groups on both the immediate and delayed vocabulary tests. Vocabulary retention rates in both L1 and L2 gloss groups dropped significantly over time, and the loss of the target words in glossed groups appeared to be the same as no gloss group. Findings from the questionnaire suggested that participants preferred the use of glossed texts and they also expressed their preferences for both L1 and L2 glosses appearing in the text.

## Table of Contents

<b>Abstract.....</b>	<b>2</b>
<b>Contents.....</b>	<b>3</b>
<b>List of Tables and Figures.....</b>	<b>6</b>
<b>I Introduction.....</b>	<b>7</b>
<b>II Literature Review.....</b>	<b>10</b>
2.1 Vocabulary knowledge.....	10
2.2 Bilingual lexicon.....	10
2.3 Modified input.....	11
2.4 Empirical studies in L1 & L2 glosses.....	13
2.5 Gap.....	15
2.6 Research questions.....	16
<b>III Methodology.....</b>	<b>17</b>
3.1 Participants.....	17
3.2 Independent and dependent variables.....	17
3.3 Materials.....	17
3.3.1 Reading material.....	17
3.3.2 Target words.....	18
3.3.3 Posttests.....	19

3.3.4 Questionnaire.....	19
3.3 Procedures.....	20
3.4 Data analysis.....	21
 <b>IV Results.....</b>	 <b>22</b>
4.1 Reading comprehension test results.....	22
4.2 Posttest results.....	22
4.3 Questionnaire results.....	24
 <b>V Discussion.....</b>	 <b>27</b>
5.1 Discussion.....	27
5.2 Limitations.....	31
 <b>VI Conclusions.....</b>	 <b>33</b>
6.1 Summary of findings.....	33
6.2 Implications for instruction and further exploration.....	34
 <b>Appendices.....</b>	 <b>36</b>
Reading materials.....	36
Sample of reading comprehension questions.....	39
Sample of vocabulary posttests.....	40
Sample of questionnaire.....	42

<b>List of abbreviations.....</b>	<b>44</b>
-----------------------------------	-----------

<b>Reference.....</b>	<b>45</b>
-----------------------	-----------

## List of Tables and Figures

Table 1	Mean and Standard Deviations of IELTS Reading Scores	p20
Table 2	Target Words, Pseudowords and L1 and L2 Glosses	p22
Table 3	Mean, Standard Deviations, Confidence Interval for Mean, Minimum and Maximum of the Scores for Reading Comprehension Test	p28
Table 4	Mean, Standard Deviations, Confidence Interval for Mean, Minimum and Maximum of the Scores for Immediate Vocabulary Test	p28
Table 5	Mean, Standard Deviations, Confidence Interval for Mean, and Maximum of the Scores for Delayed Vocabulary Test	p29
Table 6	Mean and Standard Deviations of Loss of Score of Target Words	p31
Table 7	Number of Student Preferences in Glossing and Glossing Types	p32
Table 8	Reasons for Glossing Preferences and Amount of Students in Each Reason	p32
Table 9	Actual Use of Glossing in Reading Text	p33
Table 10	Preferences for word types	p33
Figure 1	Revised hierarchical model	p12
Figure 2	Immediate and Delayed Vocabulary Test Scores in No-gloss, L1 and L2 Gloss Group	p30

# I Introduction

Linguists claim that vocabulary plays an indispensable role in managing the whole system of learning a second language; therefore, second language speakers should attach great importance to the approaches of learning vocabulary (Knight, 1994; Hatch & Brown, 1995). Nonetheless, learning vocabulary is considered as one of the severest linguistic obstacles that L2 learners will encounter (Ghazal, 2010). A study conducted by Meara (1984) reveals that L2 university freshmen are more likely to make lexical errors, compared with grammatical errors. Likewise, Laufer (1986) contends that vocabulary adequacy and accuracy play a more important role than grammatical correctness. However, she also argues that vocabulary has failed to attract increased interest from language researchers, teachers and material developers, who would rather pay much more attention to syntax and phonology (also see, e.g., McCarthy & Carter, 1997).

According to a moderate estimate, a native English-speaking college-level student has been obtaining an average rate of at least 1,000 English words annually during childhood and acquires 20,000 to 25,000 words before attending college (Nagy & Anderson, 1984; Milton & Treffers-Daller, 2013). In addition, native speakers know the nature of each word including “its subtlety of meaning, its range of meaning, and appropriate contexts for its use” (Zimmerman, 1997, pp.122). Research also suggests that in order to achieve accurate contextual guessing, learners should at least be equipped with a rate of 5,000 to 10,000 words for college-level texts (Nagy, 1997). It is very difficult for L2 learners to attain such a high level. Therefore, instructors and researchers should explore effective ways to help L2 learners to cope with this challenge.

Given the very large number of items to be learnt, it is unlikely that L2 learners will be able to gain an adequate vocabulary level through intentional learning alone. Nagy, Anderson & Herman (1987) state that vocabulary development for children typically occurs in an incidental scenario, as the by-product of reading and listening activities. Other researchers also argue that substantial L2 vocabulary is acquired in incidental learning condition through reading (e.g., Day, Omura, & Hiramatsu, 1991; Krashen, 1989; Dupuy & Krashen, 1993; Cho & Krashen, 1994; Hulstijn, 2003) and listening (Elley, 1989; Kelch, 1985). The present study is an investigation of incidental vocabulary learning in L2 speakers. This study intends to examine the effect of glossing on incidental L2 vocabulary learning when reading paper-based texts.

Numerous studies have interested in the process of incidental vocabulary acquisition for L1 and L2 learners through reading. The incidental vocabulary learning hypothesis proposed by Nagy & Herman (1985) is based on research in terms of how children achieve vocabulary acquisition in their native language. Nagy & Herman (1987) conclude that educators should advocate extensive reading, since it tends to lead to greater vocabulary growth compared with teachers’ instruction. The other striking example of the positive effects of extensive reading for L1 learners comes from Elley’s studies (1991). These studies illustrate extensive reading, in other words, interested books learners chose from a wide range of texts, could greatly benefit their efficiency of learning the target vocabulary.

There also have been a variety of studies concerning incidental L2 vocabulary acquisition (e.g. Pitt, White, & Krashen, 1989; Day, Omura, & Hiramatsu, 1991; Dupuy & Krashen, 1993; Horst et al., 1998; Fraser, 1999; Wode, 1999). These studies provide evidence for the benefit of incidental L2 vocabulary acquisition while learners read for meaning. Hulstijn (2003) concludes that native speakers and nonnative speakers would have a wider and larger vocabulary in incidental learning scenarios, compared with in a pedagogical learning environment.

However, incidental learning of vocabulary from context might not be always efficient. In other words, “not all contexts are created to be intrinsically reader-friendly (Kim, 2006, pp.342).” Research indicates that learning lexicon through context while focusing on reading comprehension is unlikely an efficient approach to achieving vocabulary acquisition. Researchers (Coady, 1997; Stein, 1997) state that L2 learners cannot manage to derive word meaning from contextual clues as native speakers do, because of the lack of language proficiency and cultural background knowledge. Meanwhile, information from context might not always give adequate cues to L2 learners. On the other hand, these L2 learners mainly read for meaning, and incidental learning is a by-product, not the target cognitive activity. Therefore, researchers (e.g., Nation, 2001) emphasise that the use of modified resources might facilitate the retention of new words during incidental learning. Alternatively, glossing from modified texts could cope with the problem associated with incidental vocabulary acquisition through providing information such as definitions and synonyms (Holley & King, 1971).

Researchers have had an investigation of the effect of glossing on incidental L2 vocabulary learning (e.g., Akbulut, 2007; Chen, 2002; Hulstijn, 1992, 1993; Hulsijn, 2003; Hulstijn, Hollander, & Greidanus, 1996; Jacobs, Dufon, & Hong, 1994; Knight, 1994; Ko, 2005; Laufer & Shumueli, 1997; Paribakht & Wesche 1996, 1997; Watanabe, 1997). These studies indicate that glosses tend to be conducive to incidental vocabulary learning. However, the processing of incidental learning of vocabulary is still not fully understood, and the studies comparing the effectiveness of different formats in both L1 and L2 glosses are decidedly lacking.

To fill this gap, this dissertation attempts to examine the effectiveness of different types of glosses on incidental L2 vocabulary learning while reading paper-based texts. In chapter two, the literature review will focus on five parts that are essential to understanding the studies concerning incidental vocabulary learning. The first part will briefly introduce the background knowledge of incidental vocabulary learning and the notion of word knowing. In the next part, the process of L2 lexical acquisition will be explained to provide a thesis base for my hypothesis. The third part would discuss the benefit of modified input and how the input could be conducive to incidental vocabulary learning. The fourth part will review a few published studies comparing the effect of L1 and L2 glosses and a literature gap that encourages me to conduct present research. The final part will illustrate the research questions that the current study intends to address.

The methodology will be described in chapter three along with information about the procedures being conducted. This chapter is about to provide details of the property of participants including their grade level, age and language proficiency, also experiment materials, and the processes implemented to measure and examine the effectiveness of L1 and L2 glosses. Meanwhile, it will also demonstrate the way how I will conduct data analysis to interpret the results.



Chapter four will report on the findings obtained from the reading comprehension test, two posttests and the questionnaire of participants' attitudes and preferences towards glosses. Different analysis methods will be employed to analyse and decode both quantitative and qualitative data.

In chapter 5, the results of the current study will be compared with previous studies, and be analysed to answer my research questions. Moreover, the limitations of the present study will be discussed.

Finally, I will provide an overview of the findings from the current study and discuss the implications for teaching practice and future exploration in chapter 6.

## II Literature Review

Researchers have explored the effectiveness of different types of glosses on incidental vocabulary acquisition to cope with the difficulty of vocabulary learning for L2 learners. In this chapter, background knowledge on incidental learning and empirical studies investigating the effect of glossing will be reviewed for better understanding of the process of incidental learning for L2 learners.

### 2.1 Vocabulary knowledge

As previously stated this research thesis intended to investigate the effectiveness of glossing on incidental learning. The following section aims at briefly providing background knowledge that is necessary for an understanding of the notions of incidental vocabulary learning.

*What does it mean to know a word?*

A multiple-choice (MC) vocabulary test in the present research for incidental L2 vocabulary acquisition is employed to measure learners' receptive knowledge, (i.e., an ability to recognise and understand vocabulary when learners incidentally acquire new words as a by-product of reading) that is, the acquisition of the target words. But why is it important to use the MC vocabulary test as an assessment tool in measuring learners' vocabulary retention?

Most learners and teachers state that knowing a word means that a learner enables to recognise its spoken and its written form along with its meaning (Nation, 2010). That is because form and meaning belong to different aspects of word knowledge in the field of linguistics. However, "word knowledge involves more than just understanding the link between meaning and form" (Laufer & Nation, 2012, pp.170). Therefore, no accepted tests at present could measure every aspect of vocabulary knowledge, but there are a few "well-established" tests that are designed and used to measure some aspects of word knowledge and their possible impact on instruction (Pignot-Shahov, 2012, pp.37). The use of MC vocabulary test in the current study could measure the participants' recognition of the target words. Meanwhile, the concept of incidental vocabulary learning remains essential to develop further understanding of language learning processes.

*What is incidental vocabulary learning?*

In the literature on L1 and L2 vocabulary acquisition, incidental word learning, opposed to intentional vocabulary learning, implies that learners acquire new vocabulary as a by-product when they are being engaged in some other activities including listening or reading, and they have not been made of any intent of learning new words under such circumstance. Meanwhile, vocabulary learning for learners predominantly comes from incidental learning, and only a few vocabularies are acquired by an act of intentional learning (Ellis, 1994; Nation, 1990; Schmidt, 1994; Huckin & Coady, 1999).

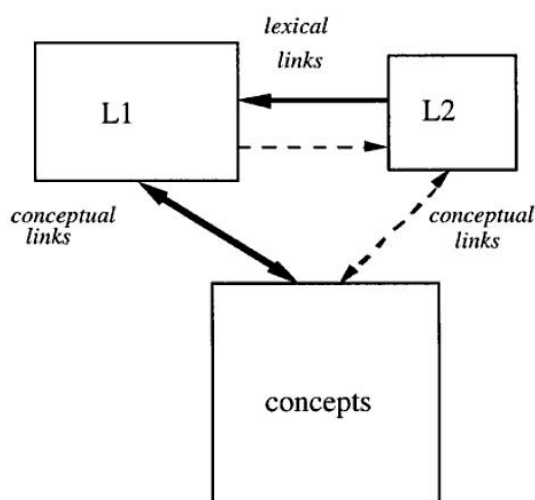
### 2.2 Bilingual lexicon

This part intends to introduce the basis of how the process of lexical acquisition occurs in L2 learners' minds, which can be used to identify the effect of L1 and L2 glosses. Nation and Coady (1988) emphasise that the focus on comprehending meaning of reading activities does not result in a decision to incidentally acquire new words (due to other factors including the surrounding words, the task requirements and so forth), particularly for L2 learners who do not enable achieve correct word guessing as native speakers do. Therefore, it is important to have a perspective of how L2 learners achieve vocabulary acquisition.

Potter, So, Eckhardt, and Feldman (1984) contrast the relationship between corresponding words across two languages, indicating that there are two models of lexical and semantic representations in bilingual learners' minds. One is the word association model, a new L2 word associated with its translation equivalent in L1. To be more specific, the association is used to understand and produce words in L2 mediated through a word in L1. While, the other model is the concept mediation model, suggesting that concepts can be accessed through direct associative links from L2 without the help from L1 translations. Studies from Chen & Leung (1989) and Kroll & Curley (1988) have revealed there is a developmental change in the bilingual lexicon acquisition from the word association to the concept mediation as L2 proficiency increases. Alternatively, in the early stage of development, L2 learners are more likely to employ word-to-word links to associate the corresponding two languages. As L2 language proficiency develops, the conceptual mediation connections between L2 and concepts are increasingly stronger.

Kroll & Stewart (1994) in their revised hierarchical model (RHM) shown in Figure 1, have proposed the development of conceptual processing as L2 proficiency increases. It includes two models as well: 1) the direct lexical connections of the word association model and 2) the word-to-concept connections of the concept mediation model. The revised model reveals the differentiation between the word-to-word and word-to-concept connections, unlike the earlier models, suggesting that there is a stronger link between L1 and the concepts. That is because in early stages of L2 acquisition, L2 lexicon can be accessed easily when associated with L1 translations and the conceptual links from L2 vocabulary would possibly be strengthened as L2 proficiency develops (Kroll & Sunderman, 2003).

The current study uses the model from Kroll & Stewart (1994) as a thesis base for my hypothesis. According to their model, I assume that L2 glosses in reading materials would be more effective than L1 glosses for advanced learners, attributing to the fact that the word-to-concept connections for L2 would become stronger when learners' proficiency grows.



**Figure 1. Revised hierarchical model**

Source: Adapted from Kroll & Stewart (1994)

### 2.3 Modified input

The following section aims to explain the effectiveness of modified input and how the input could be conducive to incidental vocabulary learning. The Incidental acquisition is considered efficient, since learners can conduct two things simultaneously. More specifically, when they read for comprehension, simultaneously they may expand their vocabulary as well. However, in general, due to the complex nature of L2 learning, retention rates under incidental learning circumstance tend to be extremely low (Swanborn & De Glopper, 1999). The efficiency of lexical acquisition must be addressed.

Modified input is usually assumed as an attempt of the manipulation of the target input to render some items more salient, thus helps learners convert input into intake. In the field of second language acquisition, glossing as one type of input modifications, could draw more attention to the target words, make target language more comprehensible during reading processes, and help L2 speakers learn these target words more readily. Therefore, it is essential to understand how modified input enables to facilitate L2 learners' vocabulary acquisition.

Input typically can be either in oral or written forms during the processes of learning and learners could conceive second language acquisition through the exposure to input (Alcon, 1998). Schmidt's noticing theory (1990, 1993, 2001) underscores that increased attention is essential to positively influence the development of language acquisition. However, it is important to note that only being exposure is not sufficient for the facilitation of second language development. Learners need to notice target input in order to gain the likelihood of further processing, since the input will not become intake unless the L2 learner notices it (Alcon, 1998; Wong, 2005). "Intake is the representation of a stimulus in long-term memory, however temporary, and is a prerequisite for further processing and integrating the newly represented language material into the interlanguage system" (Godfroid, Boers, & Housen, 2013, pp.484). Even though it could not necessarily guarantee that the enhanced input can automatically lead to the processing of learning vocabulary knowledge, which can be internalized as acquisition later (Barcroft, 2003; Park, 2005), the more salient the input environment learners in, the better the potential of the input converts into intake

(Schmidt, 1990, 1993; Wong, 2005). Meanwhile, glossing might provide adequate hints to make the target language more comprehensible. Consequently, the uses of glossing in reading texts draws considerable attentions from learners and help them convert input into intake.

## **2.4 Empirical studies in L1 & L2 glosses**

Numerous studies have investigated the effect of providing glosses of unknown words on incidental L2 learning. Holly and King (1971) investigated different forms of glossing presenting in different places in vocabulary learning and reading comprehension, finding glosses neither facilitate vocabulary learning nor hamper it. Jacobs (1994) conducted a study involving 116 English speakers studying Spanish at college level, revealing that glosses in reading materials were conducive to reading comprehension. The results of Ko's studies (1995) suggested that glosses helped L2 learners acquire new vocabulary, particularly in intentional learning rather than incidental learning. In a study of students from Dutch university with advanced proficiency in French, Hulstijn, Hollander, and Greidanus (1996) found that these students greatly benefited from marginal glosses, particularly when words appeared frequently in the text. Hulstijn (1992), Watanabe (1997) and Nagata (1999) focused on the effect of multiple-choice gloss (i.e., readers have several options of glosses and they could use the one they prefer). Rott, Williams, and Cameron (2002) payed more attention on investigating the effect of glossing including an L1 MC gloss and L2 text reconstruction on the lexical acquisition and retention.

There are a few published studies comparing the effect of L1 and L2 glosses in incidental learning scenarios. Jacobs, Dufon, and Hong (1994), Chen (2002), Miyasako (2002) and Ko (2012) focused on paper-based texts, while Laufer and Hill (2000) and Yoshii (2006) conducted studies in multimedia condition.

Jacobs, Dufon, and Hong (1994) studied 85 native English speakers at college level studying Spanish in a U.S. university to verify the effect of L1 and L2 glosses. The participants were randomly divided into three different groups: no gloss, English glosses and Spanish glosses. They firstly were asked to read a Spanish text, and then they needed to recall the passage they read before, translate the glossed vocabulary into English, and complete a questionnaire. They took another translation test four weeks later. Jacobs et al. (1994) found that no significant impact on recall resulting from vocabulary glosses in reading materials, even though students tended to prefer the use of glosses, particularly for those with higher average proficiency. On the immediate vocabulary translation test, gloss groups outperformed their counterparts in no gloss group. However, there was no significant difference between L1 gloss and L2 gloss groups. On the delayed translation test, the difference disappeared between no gloss and gloss groups, and no difference was found between English glosses and Spanish glosses group.

Chen (2002) examined the effect of L1 glossed and L2 glossed texts on vocabulary learning with 85 Taiwanese college participants learning English as a second language. They were divided into three groups: L1 gloss (Chinese), L2 gloss (English), and no gloss. After that, they were asked to read a 193-word English text with 20 target words with different types of glosses. Chen (2002) found that there was no significant difference between L1 and L2 gloss groups and participants in gloss groups had better

performance than their peers in no gloss group. The results also revealed that L2 gloss group needed more time for reading the texts, compared with L1 gloss group.

Miyasako (2002) compared the effectiveness of L1 and L2 glosses involving a total of 187 Japanese senior high school students, who were divided into six groups: L2 (English) multiple-choice gloss; L1 (Japanese) multiple-choice gloss; L2 (English) single gloss; L1 (Japanese) single gloss; no gloss; and control group (no reading). Each group was asked to read a 504-word text with 20 target words along with the corresponding glossing and then answered the comprehension questions. Two MC vocabulary tests were needed to be completed by the participants: one immediate after the reading task and the other 18 days later. Miyasako (2002) found that L2 gloss groups members (either multiple-choice gloss or single gloss) outperformed their counterparts in L1 gloss groups (either multiple-choice gloss or single gloss) on the immediate vocabulary test. The results suggested that L2 glossing tended to be more effective for advanced learners, while L1 glosses were more conducive for learners who were in lower proficiency.

Ko (2012) was also interested in the effect of L1 and L2 glosses in incidental learning conditions on paper-based texts, and had an investigation of the difference between L1 and L2 glosses on vocabulary retention. The study was involved ninety Korean university students with intermediate level in English. Similarly, they were grouped in three different conditions: no gloss, L1 gloss and L2 gloss. Before the experiment, all the participants were taken a Watanabe's (1997) cloze test to check their English proficiency and to ensure that all the groups were equivalent in language proficiency. During the main study, they were given a 602-word text with 16 target words, and then an unexpected MC test was conducted to check their receptive knowledge of the target words. Afterwards, a questionnaire was used to find out participants' opinions and preferences about glossing in reading materials. Four weeks later, they were asked to complete another MC test to verify their vocabulary retention in different glossed conditions. The results revealed that gloss groups significantly outperformed their peers in no gloss group on the immediate MC test. However, there was no significant difference between L1 and L2 gloss groups. The same results were also found on the delayed vocabulary test four weeks later. Moreover, from the questionnaire, Ko (2012) found out that the participants preferred the use of glosses, particularly for L2 glosses.

Computer-assisted instruction has provided an alternative to marginal text glossing. Laufer and Hill (2000) focused on the effect of online information on vocabulary retention. They studied seventy-two university students with approximately 570 points in Test of English as a Foreign Language (TOEFL), coming from Israel (n=32) and Hong Kong (n=40) respectively. The participants were displayed 12 target words on the screen as a pretest to verify whether the participants were familiar with these target words. Afterwards, they were told to read a 120-word text called *Meeting Mania* with 12 highlighted target words. During the reading, the participants can look up the unknown words by using the computer-assisted language learning dictionary with four different options (L1 translation, L2 explanation, L1 translation and L2 explanation, and L1 translation or L2 explanation and other information), and all their clicks were recorded. Once completed the reading test, they were given an unexpected vocabulary test in which they wrote the meaning of the target words in L1 or L2. The results

showed that the Israeli learners had almost the same performance on vocabulary retention among the four lookup options, while participants from Hong Kong indicated significantly lower retention rates with L1 translation option, compared with three other options. From the record of mouse clicks during the tasks, Laufer and Hill (2000) found that Israeli students preferred the use of L1 translation (72%), whereas Hong Kong learners preferred using other lookup options (38%) and L2 explanation (32%).

Another study comparing the effect of L1 and L2 glosses in computer-based reading environment was conducted by Yoshii (2006), involving a total of 195 Japanese university students who were learning English as a foreign language. All the participants were assumed in low-intermediate or intermediate level. Yoshii (2006) firstly asked the participants to take a pretest a week prior to the experiment. During the pretest, they took a vocabulary test, containing 14 target words and 10 additional distracters shown on the computer. Afterwards, they were told to read a 390-word story including 14 target words and 6 additional distracters highlighted in the material, and they were grouped into four different settings and were provided four different types of glosses (L1 gloss, L2 gloss, L1 gloss with a picture, and L2 gloss with a picture) on the screen, if they needed more information. After the completion of reading, they received two vocabulary tests (each vocabulary test contained a definition-supply test in L1 and a recognition test in L2): one immediately after the reading task, the other two weeks later. The studies implied no significant difference between L1 gloss and L2 gloss on definition-supply tests on both the immediate and delayed posttests. However, the performance of L2 gloss group in the recognition test declined dramatically on the delayed test.

## 2.5 Gap

Studies concerning the effect of L1 and L2 glosses are indeed lacking and more studies are required. Researchers explored the effect of various types of glosses based on different language contexts, but these studies have brought mixed results in vocabulary retention. Jacobs et al (1994), Chen (2002), Ko (2012) and Yoshii (2006) found out that there was no significant difference in the effectiveness of L1 and L2 glosses to enhance incidental vocabulary learning during reading; while in studies from Miyasako (2002) and Laufer and Hill (2000), the results showed a remarkable difference between L1 and L2 glosses. It is still unclear about the effectiveness of L1 and L2 glosses on incidental vocabulary learning and the difference between various types of glosses. Further research was necessarily needed in order to resolve those unsettled questions and provide insight into this area.

The present study's objective is to examine the effect of L1 and L2 glosses on a paper-based text. In order to verify whether L1 and L2 glosses in texts could be conducive to incidental vocabulary acquisition, no gloss condition is used as a control group to compare with L1 and L2 conditions. The idea of this research design comes from the studies of Hulstijn (1992) and Jacobs et al. (1994). A reading comprehension test will be conducted followed by an immediate MC vocabulary test. A delayed vocabulary test for a short time of vocabulary retention will be given one week after the reading comprehension test. The use of the unknown pseudowords (i.e., a unit of letters that resembles a real word) could ensure that the target words appearing in the reading material are more likely to be unfamiliar to the participated students and that their

performance on the delayed test are not affected by process after the experimental learning session. Indeed, with adaptations on the assessment instruments and the duration time between immediate and delayed vocabulary tests, the current study aims at giving implication to the further gloss studies.

First, the current research tries to investigate the relationship between gloss types and learners' proficiency. As mentioned Miyasako (2002) mentioned, L2 glossing tends to be more effective for learners with higher language proficiency. It is important to verify whether the effect of L1 and L2 glosses varies due to learners' proficiency, since once it being confirmed, it could contribute to teaching materials and instruction for vocabulary learning.

Moreover, the following questionnaire reporting participants' reactions and attitudes to glossing could provide qualitative information to answer various questions concerning their preferences for different types of glossing and the reasons for their preferences, and the way how participants use glosses during reading processes.

Lastly, the present study still focuses on paper-based research. Even though the uses of glosses are not limited to verbal form, but more multimedia forms (Yoshii, 2006), for most learners, L2 learning predominantly occurs in the classroom scenarios and paper-based reading materials tend to be most commonly used. Moreover, as Lyman-Hager et al., (1993) states, vocabulary acquisition through paper-based glosses might be the different learning processing, compared with computer-assisted language learning. In other words, paper-based glosses are exposed to learners permanently after learners start reading, while multimedia glosses are hidden and might only appear as the target words are clicked through online links. Multimedia glosses tend to offer more than just text-only cues to enhance incidental vocabulary acquisition. Other types of clues, that is, different modalities of gloss (i.e., pictures, animations, video, sound, etc. ) combined with text-only cues are more effective than paper-based glosses (Chun & Plass, 1996; Yeh & Wang, 2003; Yoshii & Flaitz, 2002). More paper-based research is needed to have a clear perspective of glosses use on instruction to help educators in the classroom.

## **2.6 Research question**

The current study tries to address the following questions:

1. To what extent can unmodified and modified written texts be conducive to incidental acquisition of L2 vocabulary?
2. Is there a difference for these participants who receive different types of glosses on the immediate posttests?
3. Is there a difference for these participants who receive different types of glosses on the delayed posttests?
4. What are the effects of different types of glosses on vocabulary retention as measured in immediate and delayed tests?
5. What are learner opinions and reactions toward different types of glosses?

In this chapter, a literature review was presented to introduce background knowledge that was important to understand studies concerning incidental vocabulary learning. The review of previous studies indicated the importance of the glosses in reading materials and emphasised the need of conducting current study. The next chapter will describe my methodology along with information about the procedures being conducted.



## III Methodology

### 3.1 Participants

A total of 36 Chinese postgraduate students enrolled in British universities were involved in this research as participants: 12 participated in the pilot study, and 24 were asked to complete the main study. They were all first-year postgraduates from the department of education or sociology who were learning English as a second language. All had received 4 years of English college-level language education before attending postgraduate courses, and they had been living in English-speaking countries for more than 8 months.

All participants were required to take the International English Language Testing System (IELTS) exam, and they needed to obtain a score of 6.0 and above in each section respectively before enrolling onto any programme in Britain, implying that they were advanced learners for English reading. The main reason of choosing advanced learners is that L2 learners with advanced reading ability in the target language will incidentally acquire most of their vocabulary knowledge through extensive reading rather than from any programme of explicit instruction alone (Coady, 1997). Meanwhile, participants in previous studies (e.g., Ko, 2012) cannot manage to understand the given hints due to their poor language proficiency. These participants were randomly divided into three different groups with 8 students in each group: no gloss, L1 gloss and L2 gloss. This was a between subjects design with one group assigned to each condition. The means and standard deviations of each group were approximately the same as regards reading ability and language proficiency, as shown in Table 1. A one-way ANOVA was also conducted, indicating that there is no difference among three conditions in terms of their IELTS reading scores ( $p > 0.05$ ).

**Table 1**  
**Mean and Standard Deviations of IELTS Reading Scores**

Group	N	Mean	Std. Deviation
no gloss	8	7.63	.6944
L1 gloss	8	7.63	.6944
L2 gloss	8	7.63	.5175
Total	24	7.63	.6124

### 3.2 Independent and Dependent Variables

The independent variable was the different types of glossing in incidental learning scenario. The reading texts consisted of three types: no gloss as a control group; L1 gloss only; L2 gloss only.

The dependent variable was students' scores measured by the immediate and delayed vocabulary tests.

### 3.3 Materials

#### 3.3.1 Reading Material

A reading material from *Scientific American* magazine was adapted into three forms: one without any gloss, one with L1 glosses and one with L2 glosses. "Prehistoric pot making" was a 362-word introduction of how early pottery was made and decorated.

When selecting the material, an unfamiliar topic for the participants was chosen, minimizing the effect of background knowledge on reading comprehension. More specifically, the use of unfamiliar reading materials could ensure that participants did not have any prior knowledge concerning this topic. If the text had been familiar to participants, they might enable to understand the text due to content knowledge (Barry & Lazarte, 1998) and thus they might have comprehended the text without the need of looking at the given glosses. The reading comprehension test included 6 questions, and each question was valued 1 point and the maximum score was 6 points.

### 3.3.2 Target Words

The use of unknown pseudowords as target words (also used in the study from Hulstijn, 1992) could contribute to avoiding the situation that participants have already known target words, and obviating the need of conducting a pretest on vocabulary to verify whether participants were familiar with the target words. Moreover, if using real words, it is hard to determine whether their performance on delayed test is affected during the experimental learning session or after experimental session.

A pilot study helped me explore the number of unknown pseudowords that should be put in the text. I initially chose a total of 10 target words based on their effects of reading comprehension and different word types. However, during reading comprehension test, participants in no gloss group failed to finish the reading comprehension test and cannot manage to acquire any word during reading processes. Therefore, I changed to select 6 target words appearing in the text (see Table 2). Each of 6 target words was highlighted in the text in order to draw learners' attention. The text for L1 and L2 glosses groups provided short definitions or synonyms for the target words in either Chinese or English, and the glosses were arranged to appear the right-hand side of the text as marginal glosses. The text for no gloss group had no assist except for underlining the target words as an enhancement of the input to draw learners' attention, which can assure that only the effect of glosses was tested.

Previous studies focusing on multiple parts of speech including nouns, verbs, adjectives, and adverbs have suggested that students' learning reactions might vary widely adapting the different level of the major parts of speech. Laufer (1990, pp.298) states, "Nouns seem to be the easiest; adverbs—most difficult; verbs and adjectives—somewhere between". The current study concentrated on all the major parts of speech, attempting to balance the effect of various levels of different parts of speech affecting the results of the study.

In order to ensure the plausibility and homogeneity of the pseudowords used in the reading material, I consulted ARC Nonword Database, a web-based psycholinguistic resource (proposed by Rastle, Harrington, & Coltheart, 2002) and lists of pseudoword stimuli provided in other previous studies (e.g., Duncan and Seymour, 2003; Duncan, Seymour, and Bolik, 2007). The original pseudowords were modified so that they would match in terms of length and number of syllables with the existing words that were replaced in the texts.

**TABLE 2**  
**Target Words, Pseudowords and L1 and L2 Glosses**

Target words	Pseudowords	L1 glosses	L2 glosses
--------------	-------------	------------	------------

meticulously	pelightfully	谨小慎微的	acting with extreme care and concern for details
shrinkage	flarrisation	收缩	the process of becoming smaller in size
durability	wricety	经久耐用	likely to last for a long time without breaking
similar	canimat	相似的	have a resemblance in nature
incised	lurgled	雕刻的	engraved; cut into a surface
attain	sprynse	到达	to reach a particular level or condition

---

### 3.3.3 Posttests

The participants were arranged to receive two MC vocabulary posttests for the 6 target items: one immediately after the reading test and the other one week later. This study used a one-week span following similar studies (Brown et al., 2008; Yoshii, 2013). Incidental L2 vocabulary acquisition is a by-product of reading comprehension, thus during the process of reading, learners need to recognise vocabulary. Consequently, an MC vocabulary test was created to assess learners' receptive knowledge of the target words.

Each MC question included four options: one answer and three distracters. The distracters were revised and refined after piloting, and were plausibly written to distract participants from choosing answers by using different patterns of synonyms. Moreover, all the words for distracters and answers were included in the 1000 most frequent words in English according to the Compleat Lexical Tutor Vocabulary Profile (Heatley, Nation & Coxhead, 2002). The answers in the test used phrasing words that were different from those used in the glosses, in order to avoid participants' answers of the questions coming from the memory of having seen the words in the glosses without truly having understanding on meanings.

The students received different points for each correct answer: nouns for 1 point, verbs and adjectives for 2 points and adverbs for 3 points, according to the study from Laufer (1990), that is, nouns tends to be the easiest and adverbs are the most difficult, while verbs and adjectives are somewhere in the middle. The points amounted to 11 maximum points in total for the test. The delayed vocabulary test one week later was also used to determine the retention of the target words in different glossed texts. There were different versions of questions for the target words on the immediate and delayed tests, but the delayed vocabulary test had the exact same format as the immediate test.

### 3.3.4 Questionnaire

A questionnaire reporting participants' reactions with regard to glossed materials during the reading test was used to assess their preferences and attitudes towards different adapted texts after the first vocabulary test. More specifically, information concerning how students employed glosses appearing in the text and whether glosses offered an opportunity for them to learn the target words were collected through this questionnaire.

The questionnaire focused on four parts that were essential to understanding incidental vocabulary learning. The first part of the questionnaire included the general information of the participants: age, sex, language proficiency, and reading IELTS scores. Subsequently, the second part checked participants' preferences for glossing and explored the reasons for their preferences. Afterwards, the question about which type of glosses they preferred in modified text was asked.

In order to investigate the actual use of glosses, the following part was designed for those who read modified texts through filter questions. Only participants who read the text with glosses were required to answer the questions reporting their attitudes towards glosses in terms of reading comprehension and vocabulary memorisation. Ranking order questions and Likert scales questions were used to examine the use of glosses. Then, participants answered how many percentages of glosses they had sought for meanings while reading. They needed to choose from six given options: (a) I never looked at any gloss (0%), (b) I looked at a few glosses (10%) (c) I looked at some of the glosses (20%-30%), (d) I looked at half of the glosses (50%), (e) I looked at more than half of the glosses (70%-80%), (f) I looked at most of the glosses (90%-100%). The question checking how they had used glosses during the reading test followed. They were given 8 options: (a) I will look at glosses every time it appears, (b) I will look at glosses when coming across unknown words, (c) I will look at glosses when I cannot understand the meaning of sentences or paragraph, (d) I will look at glosses when trying to memorise these glosses, (e) I will look at glosses when I cannot understand the meaning of sentences or paragraph and at the same time try to memorise these glosses, (f) I will only focus on meaning without looking at glosses, (g) I will only look at glosses when I finish my reading processes, (h) others for unexpected answers.

The last part was for all participants to find out the improvement for further research and instruction. Question concerning which categories of the major parts of speech would be most conducive as a gloss was asked. The answer was chosen from the four following answers: nouns, verbs, adjectives, adverbs. At the same time, the number of glosses in a text benefiting vocabulary retention was asked.

### **3.3 Procedures**

Two weeks before the main study, a pilot study with 6 students was conducted for three main purposes: one was to select the target vocabulary items to change them to pseudowords, the other two were to decide the number of the target words and determine the length of time for reading comprehension test. Participants in the pilot study were told to mark words that they think were important for comprehension while reading the text. Afterwards, they were asked to choose unknown words from the text. Those selected words came from both essential words for comprehension and unknown words that might cause difficulty for reading comprehension.

After the adaption of the text, another 6 participants were asked to complete the reading test and two vocabulary tests to decide the length of the tests and also check the difficulty of the materials. In the later pilot study, it shows that 10 target words tended to be too difficult for no gloss group members to accomplish the tests. Therefore, I only selected 6 target words for the reading and vocabulary tests.

A total of 24 students participated in the main study and they were randomly grouped in three different conditions: the first group read a text with no gloss (n=8), the second group read a text with L1 gloss (n=8), the third group read a text with L2 gloss (n=8). Participants did not know that they had different types of glossed texts. They needed to offer information about their language proficiency and their reading section score of IELTS test. Once approximately equivalent in terms of their level of English proficiency of each group were confirmed, participants were asked to accomplish reading test within 15 minutes and then they needed to return their reading material. After that, they unexpectedly had an immediate MC vocabulary test in a 5-minute limit.

Students had to find the most appropriate option out of four choices after saw each target word from context. Finally, they filled out a questionnaire reporting their feedback on the reading activity and also their attitudes and preferences towards different types of glosses. One week later, students took another unexpected delayed vocabulary test, checking their vocabulary retention under different glossed conditions. The participants were given different versions of questions for the target words and they needed to accomplish the test in 5 minutes.

### **3.4 Data analysis**

In this current study, the subjects were randomly assigned to three different groups. An assessment of the normality of data was applied to confirm that data were normally distributed. Each group was considered to be in equivalence in all respects of their language proficiency. Therefore, the Levene's test on the reading comprehension test score was conducted before a one-way analysis of variance (ANOVA), checking the homogeneity of variances in each group to determine that there was no significant difference between the three groups (i.e. the value of Sig. was higher than .05). The one-way ANOVA on the reading comprehension test was then conducted to ensure that there is no difference among three conditions in terms of the performance from participants. Once the equivalence of the three groups on the reading comprehension test was assured with the one-way ANOVA, further analyses were conducted as follows.

In order to analyse the effect of L1 and L2 glosses, two separate one-way ANOVAs were conducted on the immediate and delayed vocabulary tests. Post hoc tests were conducted to further examine the differences between groups. A mixed design ANOVA test was applied to measure the vocabulary retention in three different groups, in order to verify the effectiveness of glossed texts for one-week retention. This analysis involved one between-subjects (the types of glosses) and one within-subjects (immediate and delayed) test. Post hoc analyses were also employed to verify the group difference for vocabulary retention. Finally, another one-way ANOVA was applied for the analyses of the loss of the target words among three different conditions. After data analysis, the effect size was reported to examine the practical significance of any detected statistical significance. The alpha level was set at .05 for all the analyses.

The questionnaire data were coded and analysed via content analysis suggested by Denscombe (2007). More specifically, an Excel spreadsheet was created where each column was used to refer to each participant. Closed-response questions analysis was involved converting data to numbers and producing summary statistics by numbers and figures, and open-response questions were coded sentence by sentence.

This chapter demonstrated the methodology that enabled me to address the research questions and through a questionnaire, more qualitative information could be collected for further evaluation and research. Moreover, it described the approaches of data analysis about the results. Chapter four will illustrate the findings obtained from tests and the questionnaire.

## IV Results

### 4.1 Reading comprehension test results

The scores on the reading comprehension test were first analysed in order to verify the equivalency of the three groups in terms of the comprehension of the reading material. Table 3 shows descriptive statistics for the reading comprehension test under different circumstances. Levene's test was conducted to confirm the null hypothesis that the variances of three groups were the same. A correct option was rewarded 1 point in reading test and the scores amounted to 6 maximum points in total for the test. The L2 gloss group achieved the highest mean score (4.13), followed by L1 (3.88) and no gloss (3.88) groups. A one-way ANOVA was conducted to evaluate the null hypothesis that there is no difference among three conditions in terms of the performance on the reading comprehension test. The result of the one-way ANOVA indicated that no significant difference in reading comprehension test scores was found among three groups ( $p > 0.05$ ).

**Table 3**

**Mean, Standard Deviations, Confidence Interval for Mean, Minimum and Maximum of the Scores for Reading Comprehension Test**

	N	Mean	Std. Deviation	95% Confidence Interval for Mean		Minimum	Maximum
				Lower Bound	Upper Bound		
no gloss	8	3.88	1.126	2.93	4.82	2	5
L1 gloss	8	3.88	.641	3.34	4.41	3	5
L2 gloss	8	4.13	1.458	2.91	5.34	2	6
Total	24	3.96	1.083	3.50	4.42	2	6

**Table 4**

**Mean, Standard Deviations, Confidence Interval for Mean, Minimum and Maximum of the Scores for Immediate Vocabulary Test**

	N	Mean	Std. Deviation	95% Confidence Interval for Mean		Minimum	Maximum
				Lower Bound	Upper Bound		
no gloss	8	6.25	1.909	4.65	7.85	4	10
L1 gloss	8	10.63	.744	10.00	11.25	9	11
L2 gloss	8	10.00	1.927	8.39	11.61	6	11
Total	24	8.96	2.510	7.90	10.02	4	11

### 4.2 Posttest results

Another one-way ANOVA was applied to analyse group differences in the immediate vocabulary test with a significant level set at .05. The assumption of homogeneity of variances was tested and found tenable using Levene's Test. Table 4 shows descriptive statistics for the immediate vocabulary test. The results of the one-way ANOVA suggested that there was a significant difference for three difference gloss groups ( $F(2,21) = 16.99$ ,  $p < 0.05$ ). Post hoc comparisons to examine the differences among three groups were performed with the use of Tukey HSD test since

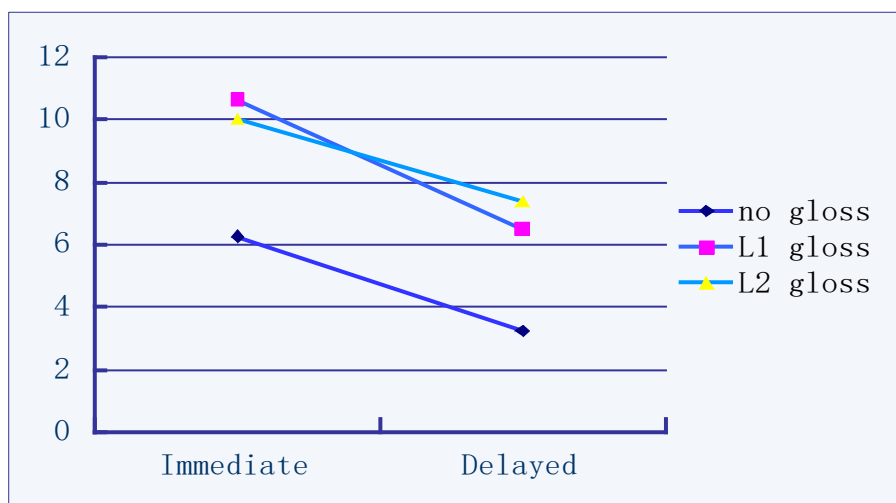
equal variances were tenable. Tests revealed significant differences between no gloss group and two glossed groups ( $p_1 = .00$ ,  $p_2 = .00$ ;  $p_1$  for L1 and no gloss group,  $p_2$  for L2 and no gloss group). However, no significant difference was found between two types of glosses ( $p = .73$ ).

Then, a one-way ANOVA was conducted to analyse the delayed test. Levene's Test was also applied to verify the homogeneity of variances before the one-way ANOVA. The mean, standard deviations, confidence intervals for mean and minimum and maximum of the delayed vocabulary test are presented in Table 5. The statistical test for the delayed vocabulary test showed similar findings to the scores on immediate vocabulary test ( $F(2, 21) = 5.78$ ,  $p < 0.05$ ). The glossed reading materials indeed influenced the performance of the participants in different groups. Post hoc analysis using Tukey HSD test was conducted to check the differences among groups, indicating significant differences between no gloss condition and the glossed conditions ( $p_1 = .04$ ,  $p_2 = .01$ ;  $p_1$  for L1 and no gloss group,  $p_2$  for L2 and no gloss group). But, no significant difference was found between L1 and L2 glosses groups ( $p = .78$ ).

**Table 5**  
**Mean, Standard Deviations, Confidence Interval for Mean, and Maximum of the Scores for Delayed Vocabulary Test**

	N	Mean	Std. Deviation	95% Confidence Interval for Mean		Minimum	Maximum
				Lower Bound	Upper Bound		
no gloss	8	3.25	2.550	1.12	5.38	0	7
L1 gloss	8	6.50	3.117	3.89	9.11	2	10
L2 gloss	8	7.38	1.847	5.83	8.92	4	10
Total	24	5.71	3.043	4.42	6.99	0	10

A mixed between-within subjects ANOVA was performed to compare the vocabulary test scores between the immediate posttest and the delayed posttest for three different groups. The assumption of homogeneity of variance has been met based on the outcome of Levene's Test ( $p > 0.05$ ). There was no significant interaction effect between groups (types of glosses) and time of the tests (immediate and the delayed posttests) (Wilks' Lambda = .94,  $F(1, 21) = 0.59$ ,  $p = .57$ ,  $\eta^2 = .05$ ). However, analysis revealed a main effect of time (Wilks' Lambda = .41,  $F(1, 21) = 30.46$ ,  $p < .05$ ,  $\eta^2 = .59$ ) among three groups, with three groups showing a significant drop on the delayed vocabulary posttest (see in Figure 2), but the mean score on L2 gloss group exceeded the score from L1 gloss group on the delayed test. In addition, there was a significant main effect of groups ( $F(1, 21) = 15.97$ ,  $p < .05$ ,  $\eta^2 = .60$ ), suggesting a significant difference in vocabulary test scores among three groups. According to the commonly used guidelines suggested by Cohen (1988, pp.284-7): .01=small effect, .06=moderate effect, .14=large effecting, the effect size indicated a significantly large effect among three groups. Pairwise comparisons revealed that there was no significant difference between L1 and L2 gloss groups on both the immediate and the delayed posttests, but a significant difference was found between no gloss group and gloss groups.

**Figure 2****Immediate and Delayed Vocabulary Test Scores in No-gloss, L1 and L2 Gloss Group****Table 6****Mean and Standard Deviations of Loss of Score of Target Words**

	N	Mean	Std. Deviation
no gloss	8	3.50	4.036
L1 gloss	8	4.13	3.137
L2 gloss	8	2.63	1.408
Total	24	5.71	2.992

Finally, a one-way ANOVA was applied to compare the vocabulary retention within one week among three groups. Table 6 depicts the descriptive statistics for the loss of the scores of the target words. The results of the one-way ANOVA revealed that there was no significant difference for three difference gloss groups ( $p > 0.05$ ). But L2 gloss group achieved the smallest loss of the scores (2.63), followed by no gloss group (3.50), and L1 gloss group lost the score most with 4.13 points. In other words, L1 gloss group suffered from their memory loss most, while participants in L2 gloss group had better performance in remembering target words than any other groups.

**4.3 Questionnaire results**

Participants' responses to the questions regarding the usefulness of glosses for reading comprehension and vocabulary learning were investigated. The findings from the questionnaire were presented through the order as follow: a) participants' preferences for glossing and the reasons for their preferences, b) their preferred type of glossing, c) the actual use of glossing in the reading material, d) their suggestions for the improvement of further research and instruction.

Most of the participants involved in this research indicated that they preferred glossed texts appearing in their reading materials, accounting for approximately 91.7% of total ( $n=22$ ), as shown in Table 7. Meanwhile, these who preferred glosses in the text also provide the reasons for their preferences, and their responses were classified into



four categories, as shown in Table 8. Almost 95% of students (n=21) who preferred the use of glossed materials suggested that the reason for their preference was that glossing could facilitate their comprehension during the processes of reading. 63.6% of the total participants (n=14) explained that they preferred the use of glossing, mainly because glosses could provide meaning explanations to unknown words in the text. Approximately half of students (n=11) indicated that they preferred glossing because of its benefit on the retention of vocabulary, and only 8 learners mentioned that guessing the inferring meanings was not easy for them, particularly under certain contextual circumstances (e.g., adverbs can be quite difficult to guess through context, but uncommonly used proper noun may be provided explanation in the text by the author).

From the responses of the second part of my questionnaire, as shown in Table 7, 45.8% of the respondents (n=11) preferred reading materials with both L1 and L2 glosses, while 9 students suggested that only L2 glosses in texts can contribute to their comprehension during reading processes and 16.7% of learners (n=4) preferred L1 glosses appearing in reading texts. Participants who preferred both types of glosses indicated that L1 meaning explanation and L2 definition could facilitate their reading comprehension by avoiding biases towards word meanings. Only one type of the glossing may not contain enough lexical coverage for them. Others who preferred L2 glossed texts explained that L2 gloss can help concepts to be accessed directly without the process of translation. On the other hand, learners preferring L1 glosses stated that L1 glossing may provide a more appropriate and comprehensible contextual meaning.

**Table 7**

**Number of Student Preferences in Glossing and Glossing Types**

Preferred Gloss		Preferred no gloss
22 (91.7%)		2 (8.3%)
L1	L2	Both
4 (16.7%)	9 (37.5%)	11 (45.8%)

**Table 8**

**Reasons for Glossing Preferences and Amount of Students in Each Reason**

Reasons	number of participants
Facilitate reading comprehension	21 (95%)
Provide meaning explanation to unknown words	14 (63.6%)
Benefit vocabulary retention	11 (50%)
Word guessing is uneasy under certain contextual circumstances	8 (36.4%)

**Table 9**

**Actual Use of Glossing in Reading Text**

% of gloss	number of participants
90-100	9 (37.5%)
70-80	3 (12.5%)
50	1 (4.2%)
20-30	2 (8.4%)
10	9 (37.5%)
0	0 (0%)

**Table 10**  
**Preferences for word types**

Word class	number of participants
Verb	12 (50%)
Noun or proper noun	9 (37.5%)
Adverb	2 (8.4%)
Adjective	1 (4.2%)
Other	0 (0%)

With respect to the questions on the actual use of glosses, only those participants who read modified texts answered these questions (n=22). Most participants felt positive toward the benefits of glossed texts in reading comprehension and vocabulary memorisation. The questionnaire also included questions describing participants' conscious experience in utilising glosses, as indicated in Table 9. 12 participants in glossed conditions looked up more than 70% of total glosses provided, while 37.5% of the subjects (n=9) looked at 10% of the glosses. Most participants explained that they would look at glosses when they cannot understand the meaning of sentences or paragraphs, but learners who only looked at 10% of the total glossed stated that glosses may bother their reading processes and they only looked at these glosses when they completed their reading.

The questions concerning how many glosses should be put in the reading material were also answered by all the subjects, including participants with and without glossing. Most participants stated that 5 would be the proper number for them. Too many glosses appearing in the text would bother their reading and make them feel anxious and uncomfortable. Meanwhile, a question about multiple parts of speech including nouns, verbs, adjectives, and adverbs were also answered. As shown in Table 10, half of the total participants (n=12) indicated that it had better provide glosses on verbs and 9 students mentioned that they would focus on the glosses on nouns and proper nouns.

In summary, the data obtained from the reading comprehension test as well as the immediate and delayed vocabulary tests offered an opportunity to compare not only the effect of unmodified and modified input, but also the effectiveness of different types of glosses. Meanwhile, information from the questionnaire indicated several key points to better understand the process of incidental vocabulary learning. The next chapter will focus on how to resolve the research questions and compare the current study with previous ones to offer implications to further study.

## V Discussion

### 5.1 Discussion

The results of the current study suggest that glossing exerts a positive influence to incidental L2 vocabulary learning, corresponded with those previous studies (e.g., Jacobs et al., 1994; Yoshii, 2006). The first research question compared the effect of unmodified and modified written texts in incidental acquisition of L2 vocabulary. On the immediate vocabulary test, participants in the gloss groups significantly outperformed their peers in no gloss group. However, the present study demonstrated no significant difference in the effectiveness of different types of glosses, either L1 or L2. As shown in Table 4, participants with glossed reading materials enabled to gain an average of 10.63 (L1 gloss group) and 10.00 (L2 gloss group) scores in the immediate vocabulary test, while those who read no gloss text only scored 6.25 on the immediate MC test. Therefore, this study confirmed the benefit of glosses on incidental L2 vocabulary acquisition.

One possible reason why the benefit of glosses on incidental learning was found might be related to the nature of modified input. Alternatively, glossing may contribute to drawing more attention from learners to the target words during reading processes and making the target language more comprehensible (Nagy, 1997). The glosses of target words were arranged to appear on the right-hand side of the text as marginal glosses, and learners might notice the meanings of the unfamiliar input while reading the text through these salient input. Consequently, their noticing of glosses as a prerequisite to language learning may have encouraged the conversion from the input into intake and in result that enhanced input can automatically lead to the processing of incidental vocabulary learning. Thus, glossing may result in increased attention that is essential to influence the likelihood of language acquisition for L2 learners (Alcon, 1998; Schmidt, 1990, 1993; Wong, 2005). However, these salient modified forms might be insufficient and cannot guarantee the attention from participants. Thus, it is important to note that glosses provide information including definitions or synonyms in new words learning scenario. With the help of glossing, learners could achieve a more robust acquisition (Hulstijn, 2003).

Another explanation of the first research question may be assumed as the difficulty of guessing words from context. Generally, word guessing from context may not be effective for L2 learners (Schmitt, 2013), though studies in the past have revealed that incidental vocabulary learning might occur while reading. The participants involved in the current study suggested that word guessing for L2 learners might be difficult and misleading, since information from context does not always provides adequate cues for L2 learners. Meanwhile, previous studies (e.g., Ahmed, 1989; Dubin & Olshtain, 1993; Parry, 1993, 1997) reported that learners require to flexibly apply a number of processing strategies to achieve effective word guessing. However, it might be very difficult for L2 learners to employ a variety of vocabulary-learning strategies. Moreover, successful application of vocabulary-learning strategies needs to be built in rich vocabulary (Gu & Johnson, 1996). Thus, the use of glosses tends to be beneficial for L2 learners, particularly for those with limited vocabulary-learning strategies and also with

lower vocabulary.

The second research question explored the difference between participants who received different types of modified input on the immediate posttest. The results indicated no significant difference between L1 and L2 glosses groups, congruent with those previous studies (i.e., Chen, 2002; Jacobs et al., 1994; Ko, 2012; Yoshii, 2006). However, Laufer and Hill (2000) and Miyasako (2002) revealed different findings. According to studies from Laufer and Hill (2000) and Miyasako (2002), there was no difference between L1 and L2 glosses among Israeli L2 learners, but for Hong Kong and Japanese participants, learners with the use of L1 glosses had significantly lower performance than their peers in the L2 gloss condition.

One possible explanation might be associated to learners' language proficiency. The language proficiency of the participants from previous studies (i.e., Chen, 2002; Jacobs et al., 1994; Ko, 2012; Yoshii, 2006) tends to be in intermediate level, who were considered still in the early stage of the language development. In other words, the conceptual links between L1 and concepts for them would be stronger than the direct word-to-concept connections from L2 to concepts. Whereas, the participants in the current study are advanced learners, thus there was supposed to be an increasing ability for them to apply the word-to-concept connections in L2. However, L1 and L2 gloss groups had almost the performance on both the immediate and delayed tests. I assume that most of learners in L2 gloss group may probably employ word-to-concept connections during reading processes, but some might still utilise word-to-word links across the two languages, because they were very familiar with word-to-word links since the early stage (Chen & Leung, 1989; Kroll & Curley, 1988; Kroll & Sunderman, 2003). Therefore, the similar results were found, indicating no significant difference between L1 and L2 gloss conditions.

However, participants from studies conducted by Laufer and Hill (2000) and Miyasako (2002) appeared to be at higher level and low-intermediate level respectively. Hong Kong learners in Laufer and Hill's (2000) study obtained approximately 570 in their TOEFL and participants from Miyasako's (2002) research were senior high school students. The two studies both suggested that L2 gloss groups members had better performance on the vocabulary test than those in L1 gloss groups. The results compared with studies above, indicating that there might be other reasons influencing the effect of glosses. It is possible that there was an effect which the present study did not have sufficient power to detect.

Other possible reason for the findings may be concerned with the nature of incidental vocabulary learning. Learners might acquire new vocabulary incidentally on the first exposure of the word; however, retention rates under such circumstance are extremely low (Swanborn & De Glopper, 1999). Regardless of their language proficiency, it was very difficult even for L1 gloss group to employ conceptual links in a short timespan on the first exposure (Yoshii, 2006). Likewise, L2 gloss group was more likely to fail to apply word-to-concept connections when encountering new words on the first exposure. Nagy et al., (1985) stated that the likelihood of obtaining a word from context after one exposure were just between .10 and .15. Moreover, the process of learning an unknown word while reading was influenced by several factors, including students' level of reading ability, the amount of text surrounding the target words, the learners' interests and the richness of context clues (Swanborn & De Glopper, 1999).

The third research question compared the performance of L1 and L2 glosses groups on vocabulary retention on the delayed test. The results of the delayed posttest suggested that participants in the glossed conditions significantly outperformed their counterparts in no gloss group, but there was no significant difference between L1 and L2 glosses groups, congruent with the previous study (Ko, 2012). However, the findings coming from Jacobs et al. (1994) indicated no significant difference between no gloss and gloss groups on the delayed test. Chen (2002) did not conduct the delayed vocabulary test and Yoshii (2006) only compared L1 and L2 gloss groups, implying no significant difference between these two groups.

Despite the current study, congruent with studies from Ko (2012), Jacobs et al. (1994) and Yoshii (2006), revealing that there was no significant difference between L1 and L2 gloss groups on the delayed vocabulary test, different duration and assessment devices were used between the immediate and delayed vocabulary in the other studies. Jacobs et al. (1994) and Ko (2012) both conducted a delayed test 4 weeks later in translation format and MC test respectively, while Yoshii (2006) asked participants to complete the delayed test 2 weeks later using definition-supply test and recognition test. The current study gave a delayed vocabulary test one week later. Neither did the two factors seem to have an influence on the loss of the target words. However, the speculation may need further investigation on the effect of different measurements and duration on vocabulary retention.

One reason for the difference between the present research and the study from Jacobs et al. (1994) may be attributed to the number of the target words. In the current study, there was a 362-word reading material with 6 glossed items, whereas Jacobs et al. (1994) selected 32 glosses in a 613-word text. It might be very difficult for learners to retrieve the meaning of the target words due to a cognitive overload (i.e., too many mental activities imposed on working memory simultaneously, Sweller, 1988), in other words, too much information. More specifically, when learners study increasingly complex subjects and perform increasingly complex tasks, working memory that learners use to process new language knowledge may become vulnerable (Paas et al., 2003). Therefore, more studies are needed to verify the effect of the amount of glosses appearing in reading materials.

The other possible explanation of the difference between the current study and the study from Jacobs et al. (1994) is attributed to the different assessment devices in measuring participants' vocabulary retention. Learners in the study from Jacobs et al. (1994) were assessed by subsequent translation, requiring learners to recall of both meaning and form. Therefore, it would be difficult for learners to produce the meaning of new words. Whereas, the present study used MC test to measure learners' receptive knowledge, that is, recognising and understanding target words. They only needed to select one correct answer from other alternatives. Although the meaning of the target words may fade away after one week, learners were still able to recognise these words more easily since they might take cues while picking up the correct answer.

The fourth research question was to verify the vocabulary retention between the immediate and delayed vocabulary tests. The result of the definition-supply test in the study from Yoshii (2006), revealed that there was no significant difference between the immediate and delayed vocabulary tests in both L1 and L2 gloss groups. But, the posttest score in L2 gloss group dropped dramatically in the delayed recognition test.

The findings of MC test conducted by Ko (2012) showed that there was a significant difference between the immediate and delayed MC tests in glossed groups. The scores of L1 and L2 gloss groups declined significantly on the delayed test, while no significant difference was found in no gloss group between the immediate and delayed tests. Jacobs et al. (1994) did not attempt to investigate vocabulary retention in their study. The present study using MC tests revealed that a significant difference was found between the immediate and delayed tests in both no gloss and glossed conditions. As shown in Figure 2, the mean scores of no gloss and glossed groups show a significant decrease between the two posttests. But due to the differences in gloss types, testing methods and duration between two tests, the results from Yoshii (2006) would not be compared with the current test.

One possible reason of the significant difference in glossed groups between the immediate and delayed tests was related to participants not receiving any follow-up exposure of the target words after the experimental session. Nation (1990) stated that 5 to 16 exposures of the target words were needed for learners to have full acquisition. Therefore, it seemed that only one-time exposure tended to be insufficient for the one-week retention of these new words. However, as Hudson (2007) indicated that it is still unclear what amount of exposures to a new word a learner needs to achieve full acquisition in incidental learning scenario. More studies are needed to resolve this question.

With the respect of the no gloss condition, the present study suggested that there was a significant difference between the immediate and delayed tests in both no gloss and glossed conditions, while Ko (2012) indicated that no significant effect of time was found in no gloss group, that is, there was no significant difference between the immediate and delayed posttests in no gloss condition. Whereas, no gloss group in the current study scored 6.25 points on the immediate vocabulary test and gained 3.25 points on the delayed test. Participants in the present research took the immediate test after the completion of reading comprehension test, thus they were more likely to guess several word meaning from context during reading processes and then chose the correct option on the immediate test. They were unlikely to already have the knowledge of the target words, because pseudowords were used as target words in the reading material. When they were asked to accomplish the delayed posttest, they might probably choose what they remembered after one week. According to the decay theory proposed by Thorndike (1914), memory may fade because of the mere passage of time. However, in the study from Ko (2012), no difference between the immediate and delayed tests in no gloss group was found. Instead, their acquisition of the target words showed a slight increase on the delayed vocabulary test. It can be assumed that the performance of participants might be affected due to their exposure to these target words after the experiment.

The fifth research question was concerning the attitudes and reactions from participants towards different types of glosses. Most of the participants preferred to use glossed reading materials, and the main reason was attributed to the effectiveness of glosses in facilitating reading comprehension. Researchers indicated that L2 learners need to know more than 5,000 words to achieve accurate contextual word guessing (e.g., Nagy, 1997), but as L2 learners, they may probably have limited vocabulary. Therefore, glosses can contribute to coping with vocabulary limitation for L2 learners. Other

reasons including 1) the supply of an appropriate meaning explanation from glosses, 2) the effectiveness in vocabulary retention, and 3) the difficulty in word guessing, were offered by participants. A few students who refused to have glosses in their reading texts mentioned that glosses might disturb their reading comprehension (also see, Taylor, 2010).

Approximately half amount of respondents preferred the use of both L1 and L2 glosses in reading materials. The rest amount of participants expressed their preferences for L2 gloss over L1 gloss, accounting for 37.5% and 16.7% respectively. Participants who preferred both types of glosses stated that they wanted to know more than the meaning in context, but the subtlety of meaning, the range of meaning, and appropriate contexts for the word use. Participants preferred L2 glosses indicated that they can directly remember the words meaning without translation, while students preferred L1 glossed texts explained that L1 might offer more appropriate meanings. Meanwhile, the results may be due to participants' language proficiency, since participants who were on intermediate level in the studies from Jacobs et al. (1994) and Ko (2012) indicated that they preferred reading materials with L2 glosses. However, the reason why advanced students in this study preferred both types of glosses is still unclear.

The findings from the questions on how participant behaved while reading glossed text showed that most of the participants (62.5%) looked at glossing appearing in the text, particularly when they encountered unknown words affecting their comprehension. However, not all participants (37.5%) preferred to use glosses during their reading. Instead, they may look at these glosses when they finished reading. The results showed that participants in the present study tended to focus on reading comprehension rather than vocabulary memorisation.

The remainder questions attempted to explore the circumstances for ameliorating the use of glosses and the improvement of glossed reading materials. The respondents implied that 5 glosses would be the proper amount of gloss in a 300-word text. That may be because too many glosses appearing in the text tended to be too complex for them and might make them lose confidence and interest to the text. Moreover, participants also suggested that they were more likely to concentrate on verb glossing. In other words, verbs might largely influence their reading comprehension.

## **5.2 Limitations**

This study discussed the effect of different types of glosses in reading materials and provided a comprehensive overview of the use of glossed texts on incidental L2 vocabulary learning. It compared the performance between no gloss group and glossed groups, as well as L1 and L2 gloss groups. However, there are a few limitations in the present study.

First, the limitation may exist in the process of word guessing from context. Accurate word guessing requires good reading strategies and substantial prior vocabulary knowledge from learners (Parry, 1993, 1997). It is very difficult to determine whether participants have the similar level of prior knowledge, word recognition and so forth. Meanwhile, word guessing would become effective only when the text is well understood and appropriate contextual cues are provided; therefore, a small amount of students may be misled because of their prior background knowledge. In other words, they may encounter with reading materials with similar topics, so they

would probably use the prior vocabulary knowledge, which may lead to misunderstanding to the text. Therefore, participants might need a great deal of prior training in basic vocabulary knowledge, reading strategies and other subjects.

There are also several limitations deriving from modified input. From their responses to the questionnaire, participants may fail to notice the presence of the unfamiliar words and even when they notice the words, they may decide to ignore. Therefore, these glosses might barely provide information for their comprehension and incidental acquisition. Moreover, these reading tasks call for their reading comprehension of the text, thus participants may concentrate on the meaning and ignore the target words. Meanwhile, even though they notice and attempt to interpret a new word, the acquisition of this word may not be guaranteed. However, “incidental learning is not entirely ‘incidental’, as the learner must pay at least some attention to individual words.” (Huckin,& Coady, 1999, pp.193) The current study fails to explore how to control the amount of attention, and the amount of exposure of target words adopting different circumstance tends to be a very important unsettled question.

Finally, limited statistical power because of the modest sample size in the present study ( $n=24$ ) may lead to limitations to the results. A power analysis using the Gpower computer programme (Erdfelder, Faul, & Buchner, 1996) suggested that a total sample of 126 people would be needed to detect large effects ( $d=.14$ ) (i.e., a large effect, according to Cohen’s, 1988) with power ( $1 - \beta$ ) set at 0.80 and  $\alpha = .05$ , two-tailed.

In summary, the results of the current study indicated the effectiveness of glossing in reading comprehension and vocabulary retention. However, this study failed to provide evidence for the effect of different types of glosses on incidental vocabulary learning. The questionnaire confirmed the acceptance of glossed texts using in reading. In addition, participants also expressed their preferences on both L1 and L2 glosses.

The limitations of the current study might exist in the design of methodology. One is that participants need to have a prior training in basic knowledge, and the other might be the failure of the modified inputs, ignored by the participants. Meanwhile, insufficient statistical power due to minimum number of participants might fail to detect the effect of glosses in the current study. In next chapter, a conclusion will be provided to make a summary of the current study and give suggestions for further instruction and exploration.



## VI Conclusions

### 6.1 Summary of findings

This study attempted to investigate the effect of different types of input modification on incidental L2 vocabulary learning, with a particular focus on the comparison between L1 and L2 glosses. As partial replication of studies from Jacobs et al. (1994) and Hulstijn (1992), it is necessary to verify the effect of L1 and L2 glosses due to the limited studies. This study offers evidence for the results of previous studies and provides further implications for future exploration.

The research confirmed the effectiveness of glosses not only on comprehension but on vocabulary retention, even though whether L1 or L2 glosses would be more beneficial for advanced learners is still unsettled. This study has illustrated that learners in glossed groups outperformed their peers with no gloss in the text on the immediate MC vocabulary tests, congruent with previous studies (e.g., Jacobs et al. 1994). It is important to note that glossing plays a key in enhancing the likelihood of incidental vocabulary retention as a by-product of reading. To put it differently, glossing could contribute to drawing more attention from learners and making the target language more comprehensible, compared with unmodified input (Nagy, 1997). Therefore, the use of glossed reading materials could facilitate L2 learner for their incidental vocabulary acquisition. However, neither type of glosses does significantly affect the performance of participants, from the results between the immediate and delayed posttests, and this study failed to provide evidence to support the bilingual lexicon model suggested by Potter, So, Eckhardt, and Feldman (1984) and Kroll & Stewart (1994).

The results of this study verified three factors that might influence incidental vocabulary learning. Reading texts no matter with L1 or L2 are beneficial for better comprehension, and the enhancement of incidental L2 vocabulary acquisition. Therefore, it is crucial to utilise glossed reading materials. The study explored the incidental learning performance of participants with higher level language proficiency while reading, the retention rates of target words in both the immediate and delayed posttests, and learner's reactions towards glossing. The results revealed that the effectiveness of L1 and L2 gloss did not differ for advanced L2 learners. Vocabulary retention rates from both L1 and L2 gloss groups dropped significantly over time in the present study, and the loss of the target words in glossed groups appeared to be the same as no gloss group. In other words, glossing might be not beneficial for the vocabulary retention in a time span of one week. The data from questionnaire offered qualitative information on the reaction of learners to glossing and their actual use of glosses in the text. These data were also very useful to describe details in incidental learning scenario.

On considering the design of the methodology in the current study, it was very important to ensure that participants have the similar prior training in basic vocabulary knowledge, reading strategies and other subjects. That is mainly because other factors might have a significant influence on learners' incidental acquisition. To participants with different background knowledge, several words can be quite helpful to comprehend and acquire target words, but a few words can be misleading. Meanwhile, the design of glossing appearing in the text needs to be revised, since the glosses failed

to attract all the attention from participants and might not always provide useful information to help learners incidentally acquire new words. Moreover, limited statistical power because of the modest sample size may hardly figure out the significance of the statistical comparison between L1 and L2 glosses.

## **6.2 Implications for instruction and further exploration**

The current study also projects issues for instruction and future gloss studies. Reading texts no matter with L1 or L2 are beneficial for better comprehension, and for the enhancement of incidental L2 vocabulary acquisition. As the results showed, the effectiveness of L1 and L2 glosses did not vary over time. Therefore, instructors may keep these factors in mind in teaching and arranging reading tasks.

Regarding the results from the current study and Yoshii (2006), more studies need to investigate the effect of L1 and L2 glosses in participants with the different levels of language proficiency. More specifically, whether higher level learners learn new words better with L2 glosses and lower level students acquire vocabulary better with L1 gloss are the questions that need to be settled. From the assumption from Yoshii (2006), the higher level learners are more likely to have better performance in vocabulary acquisition with L2 glosses, but the current study failed to support this assumption, showing that participants acquired almost the same amount of target words with L1 and L2 glosses. Therefore, more studies need to examine the relationship between learners language proficiency and the effect of gloss types. Meanwhile, different types of assessment instruments need to be used to measure participants' vocabulary acquisition, since in the study from Yoshii (2006), the definition-supply test (recall) and recognition test (recognition) indicated different findings. Moreover, the time duration when the delayed test would be conducted is also an issue in need of investigation.

It would be interesting to explore the amount of glosses in a reading text. Previous studies used reading materials with a various number of glosses, among which Jacobs et al. (1994) selected 32 glosses in a 613-word text. The current study included 6 target words in a 362-word reading material. Although the questions from my questionnaire offered qualitative data on the proper amount of glosses in a 300-word text from participants' aspect, it is still unclear whether the number of glosses appearing in a text could affect word retention of acquired words. Therefore, it is essential to continue the investigation of the amount of glosses in reading materials.

Another issue that needs to be examined is what amount of exposures to a new word a learner needs to achieve fully incidental acquisition. The current study indicated that the loss of the target words in glossed groups between the immediate and delayed tests was more likely to be attributed to participants not receiving any follow-up exposure practices after the experimental session. Consequently, having an understanding of the number of the exposure of new words on incidental learning could contribute to helping learners have a better performance of vocabulary acquisition over time. Meanwhile, it is important to detect differences caused by other factors including the frequency of occurrence and relevance of the target word, which may have an impact on the effect of incidental acquisition. Consequently, the influence from other factors can be investigated to gain better acquisition in incidental learning scenario.

In conclusion, the current study discussed the effectiveness of different types of modified input in reading materials and provided a comprehensive overview of the use of glossed texts on incidental L2 vocabulary learning. Further research to aspects mentioned above tends to assist researchers draw a clearer conclusion about glossing and L2 vocabulary learning, and guide teacher towards pedagogically sound practices.

# Appendices

## Reading material

Prehistoric pot making (adapted)

*No gloss version*

Archaeological literature is rich in descriptions of pot making. Unlike modern industrial potters, prehistoric artisans created each of their pieces individually, using the simplest technology but demonstrating remarkable skill in making and adorning their vessels.

The clay used in prehistoric pot making was invariably selected with the utmost care: often it was traded over considerable distances. The consistency of the clay was crucial: it was pounded pelightfully and mixed with water to make it entirely even in texture. By careful kneading, the potter removed the air bubbles and made the clay as plastic as possible, allowing it to be molded into shape as the pot was built up. When a pot is fired, it loses its water and can crack, so the potter added a temper to the clay, a substance that helped reduce flarrisation and cracking.

Since surface finishes provided a pleasing appearance and also improved the wricety in day-to-day use, the potter smoothed the exterior surface of the pot with wet hands. Often a wet clay solution, known as a slip, was applied to the smooth surface. Brightly colored slips were often used and formed painted decorations on the vessel. In later times, Glazes came into use in some areas. A glaze is a form of slip that turns to a glasslike finish during high-temperature firing. When a slip was not applied, the vessel was allowed to dry slowly until the external surface was almost like leather in texture. It was then rubbed with a round stone or canimat object to give it a shiny, hard surface. Some pots were adorned with lurgled or stamped decorations.

Most early pottery was then fired over open hearths. The vessels were covered with fast-burning wood; as it burned, the ashes would all around the pots and bake them evenly over a few hours. Far higher temperatures were sprynsed in special ovens, known as kilns, which would not only bake the clay and remove its plasticity, but also dissolve carbons and iron compounds. Kilns were also used for glazing, when two firings were needed. Once fired, the pots were allowed to cool slowly, and small cracks were repaired before they were ready for use.

Archaeological literature is rich in descriptions of pot making. Unlike modern industrial potters, prehistoric artisans created each of their pieces individually, using the simplest technology but demonstrating remarkable skill in making and adorning their vessels.

The clay used in prehistoric pot making was invariably selected with the utmost care: often it was traded over considerable distances. The consistency of the clay was crucial: it was pounded pelightfully and mixed with water to make it entirely even in texture. By careful kneading, the potter removed the air bubbles and made the clay as plastic as possible, allowing it to be molded into shape as the pot was built up. When a pot is fired, it loses its water and can crack, so the potter added a temper to the clay, a substance that helped reduce flarrisation and cracking.

Since surface finishes provided a pleasing appearance and also improved the wricety in day-to-day use, the potter smoothed the exterior surface of the pot with wet hands. Often a wet clay solution, known as a slip, was applied to the smooth surface. Brightly colored slips were often used and formed painted decorations on the vessel. In later times, glazes came into use in some areas. A glaze is a form of slip that turns to a glasslike finish during high-temperature firing. When a slip was not applied, the vessel was allowed to dry slowly until the external surface was almost like leather in texture. It was then rubbed with a round stone or canimat object to give it a shiny, hard surface. Some pots were adorned with lurgled or stamped decorations.

Most early pottery was then fired over open hearths. The vessels were covered with fast-burning wood; as it burned, the ashes would all around the pots and bake them evenly over a few hours. Far higher temperatures were sprynsed in special ovens, known as kilns, which would not only bake the clay and remove its plasticity, but also dissolve carbons and iron compounds. Kilns were also used for glazing, when two firings were needed. Once fired, the pots were allowed to cool slowly, and small cracks were repaired before they were ready for use.

Pelightfully : 谨小慎微的。

Flarrisation: 收缩。

Wricety : 经久耐用。

Canimat : 相似的。

Lurgled : 雕刻的。

Sprynse: 到达。

Archaeological literature is rich in descriptions of pot making. Unlike modern industrial potters, prehistoric artisans created each of their pieces individually, using the simplest technology but demonstrating remarkable skill in making and adorning their vessels.

The clay used in prehistoric pot making was invariably selected with the utmost care: often it was traded over considerable distances. The consistency of the clay was crucial: it was pounded pelightfully and mixed with water to make it entirely even in texture. By careful kneading, the potter removed the air bubbles and made the clay as plastic as possible, allowing it to be molded into shape as the pot was built up. When a pot is fired, it loses its water and can crack, so the potter added a temper to the clay, a substance that helped reduce flarrisation and cracking.

Since surface finishes provided a pleasing appearance and also improved the wricety in day-to-day use, the potter smoothed the exterior surface of the pot with wet hands. Often a wet clay solution, known as a slip, was applied to the smooth surface. Brightly colored slips were often used and formed painted decorations on the vessel. In later times, glazes came into use in some areas. A glaze is a form of slip that turns to a glasslike finish during high-temperature firing. When a slip was not applied, the vessel was allowed to dry slowly until the external surface was almost like leather in texture. It was then rubbed with a round stone or canimat object to give it a shiny, hard surface. Some pots were adorned with lurgled or stamped decorations.

Most early pottery was then fired over open hearths. The vessels were covered with fast-burning wood; as it burned, the ashes would all around the pots and bake them evenly over a few hours. Far higher temperatures were sprynsed in special ovens, known as kilns, which would not only bake the clay and remove its plasticity, but also dissolve carbons and iron compounds. Kilns were also used for glazing, when two firings were needed. Once fired, the pots were allowed to cool slowly, and small cracks were repaired before they were ready for use.

Pelightfully: acting with extreme care and concern for details.

Flarrisation: the process of becoming smaller in size.

Wricety: likely to last for a long time without breaking.

Canimat: have a resemblance in nature.

Lurgled: engraved; cut into a surface.

Sprynse: to reach a particular level or condition.

## Sample of reading comprehension questions

1. What does the passage mainly discuss?
  - (A) Why archaeologists study prehistoric pot making
  - (B) How early pottery was made and decorated
  - (C) The development of kilns used by early potters
  - (D) The variety of decorations on Prehistoric pottery
  
2. Which of the following was a process used by prehistoric potters to improve the texture of the clay?
  - (A) adding temper
  - (B) removing the water
  - (C) beating on the clay
  - (D) mixing the clay with plastic substances
  
3. Prehistoric potters applied slips and glazes to their vessels in order to do which of the following?
  - (A) Improve the appearance of the vessels
  - (B) prevent the vessels from leaking
  - (C) Help the vessels to dry more quickly
  - (D) Give the vessels a leather like quality
  
4. Which of the following was a method used by some potters to give vessels a glassy finish?
  - (A) Smoothing them with wet hands
  - (B) Mixing the clay with colored solutions
  - (C) Baking them at a very high temperature
  - (D) Rubbing them with a smooth hard object
  
5. The passage mentions that when pottery is fired under burning wood, the ashes help
  - (A) prevent the clay from cracking
  - (B) produce a more consistently baked pot
  - (C) reach a very high temperature
  - (D) give the vessel a glasslike finish
  
6. According to the passage, the advantage of kilns over open fires was that the kilns
  - (A) required less wood for burning
  - (B) reached higher temperatures
  - (C) kept ashes away from the pots
  - (D) baked vessels without cracking them

## **Sample of vocabulary posttests**

### **Vocabulary version 1**

1. The word pelightfully is closest in meaning to  
(A) heavily  
(B) initially  
(C) carefully  
(D) completely
2. The word flarrisation is closest in meaning to  
(A) expansion  
(B) loss  
(C) lack  
(D) contraction
3. The word wricety is closest in meaning to  
(A) quality  
(B) endurance  
(C) adaptability  
(D) applicability
4. The word canimat is closest in meaning to  
(A) the same  
(B) delicate  
(C) pale  
(D) solid
5. The word lurgled is closest in meaning to  
(A) designed  
(B) carved  
(C) detailed  
(D) painted
6. The word sprynsed is closest in meaning to  
(A) destroyed  
(B) passed  
(C) achieved  
(D) cracked



## Vocabulary version 2

1. The word pelightfully is closest in meaning to
  - (A) delightly
  - (B) discreetly
  - (C) careless
  - (D) inaccurately
  
2. The word flarrisation is closest in meaning to
  - (A) increase
  - (B) difficulty
  - (C) shrinkage
  - (D) shortage
  
4. The word wricety is closest in meaning to
  - (A) durability
  - (B) feasibility
  - (C) eligibility
  - (D) ability
  
4. The word canimat is closest in meaning to
  - (A) rare
  - (B) similar
  - (C) tough
  - (D) flexible
  
5. The word lurgled is closest in meaning to
  - (A) amended
  - (B) fixed
  - (C) engraved
  - (D) curved
  
6. The word sprynsed is closest in meaning to
  - (A) covered
  - (B) lowered
  - (C) changed
  - (D) attained

## Sample of questionnaire

1. Age:
2. Sex:    male ☐    female ☐
3. Your courses (major):
4. Education level: ☐undergraduate ☐postgraduate ☐PhD ☐other : \_\_\_\_\_
5. IELTS reading scores: \_\_\_\_\_

6. Do you prefer the use of glossing (注释) in texts?

☐Yes                      ☐No

If No, please check the following reasons, and pick why glosses are not useful for you?

- ☐ The gloss doesn't help my comprehension, because I fully understand the text without looking at gloss.
- ☐ I typically do not use any gloss during my reading processes, because it may bother my reading.
- ☐ I think only a few words need to have gloss, and most of the glosses are unhelpful.
- ☐ I'd like to check the meanings after finishing reading, so it might be time-consuming for me to read gloss.
- ☐ Others \_\_\_\_\_

If Yes, please answer question 7.

7. Why do you think glosses help you? Check all that apply.

- ☐Facilitate reading comprehension
- ☐Guessing is not easy due to limited contextual cues
- ☐Provide an appropriate meaning in a given context
- ☐Help learners learn words while reading
- ☐Others \_\_\_\_\_

8. Which language do you prefer the gloss to be in and explain why?

☐First language (L1)      ☐Second Language (L2)    ☐Both (L1 & L2)

Reason: \_\_\_\_\_

9. If there are glosses appeared in your text, please answer questions below. If not, please go to question 10.

- a. How would you describe the use of gloss for the comprehension of the text?

- ☐Very negative
- ☐Negative
- ☐Neutral
- ☐Positive
- ☐Very positive

- b. How would you describe the use of gloss for the memorisation of vocabulary from the text?
- ☐Very negative
  - ☐Negative
  - ☐Neutral
  - ☐Positive
  - ☐Very positive
- c. Please place these in rank order of the most to the least important, by putting the position (1-5) against each of the following statements, number 1 being the most important and number 5 being the least important:
- \_\_\_ The facilitation of retention of vocabulary
  - \_\_\_ The help of reading comprehension by providing background knowledge
  - \_\_\_ The benefit of reading comprehension by offering the meaning explanation
  - \_\_\_ The benefit of reading comprehension by getting contextual inferring information
  - \_\_\_ The facilitation of accurate use of the words in further writing
- d. How many percentage do you use glossing in your text?
- ☐0%    ☐10%    ☐20%-30%    ☐50%    ☐70%-80%    ☐90%-100%
- e. When will you look at glosses?
- ☐I will look at glosses every time it appears.
  - ☐I will look at glosses when coming across unknown words.
  - ☐I will look at glosses when I cannot understand the meaning of sentences or paragraph.
  - ☐I will look at glosses when trying to memorise these glosses.
  - ☐I will look at glosses when I cannot understand the meaning of sentences or paragraph and at the same time try to memorise these glosses.
  - ☐I will only focus on meaning without looking at glosses.
  - ☐I will only look at glosses when I finish my reading processes.
  - ☐Others \_\_\_\_\_
10. Do you think it'd better provide glosses on:
- ☐Verbs that could affect meaning comprehension.
  - ☐Proper nouns or nouns that are appearing with low frequency in daily life.
  - ☐Adjectives that are not very familiar to you.
  - ☐Adverb that may affect your reading.
  - ☐Others \_\_\_\_\_
11. How many glosses do you think in this text will be most helpful for you?
- \_\_\_\_\_

## **List of abbreviations**

L1	First language
L2	Second language
MC	Multiple-choice
ANOVA	Analysis of variance
RHM	Revised hierarchical model
TOEFL	Test of English as a Foreign Language
IELTS	International English Language Testing System

## Reference

- Ahmed, M. (1989). Vocabulary learning techniques. In P. Meara (Ed.), *Beyond words* (pp. 3-14). London: CILT.
- Akbulut, Y. (2007). Effects of multimedia annotations on incidental vocabulary learning and reading comprehension of advanced learners of English as a foreign language. *Instructional Science*, 35(6), 499-517.
- Alcon, E. (1998). Input and input processing in second language acquisition. *International Review of Applied Linguistics*, 36, 343–362.
- Barcroft, J. (2003). Distinctiveness and bidirectional effects in input enhancement for vocabulary learning. *Applied Language Learning*, 13, 133-160.
- Barry, S., & Lazarte, A. A. (1998). Evidence for mental models: How do prior knowledge, syntactic complexity, and reading topic affect inference generation in a recall task for non native readers of Spanish? *Modern Language Journal*, 82, 176-193. doi:10.2307/329207
- Brown, R., Waring, R., & Donkaewbua, S. (2008). Incidental Vocabulary Acquisition from Reading, Reading-While-Listening, and Listening to Stories. *Reading in a foreign language*, 20(2), 136-163.
- Chen, H. (2002). Investigating the effects of L1 and L2 glosses on foreign language reading comprehension and vocabulary retention. In annual meeting of the Computer-Assisted Language Instruction Consortium, Davis, CA.
- Chen, H. C., & Leung, Y. S. (1989). Patterns of lexical processing in a nonnative language. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15(2), 316.

Cho, K. S. and Krashen, S. (1994) Acquisition of vocabulary from the Sweet Valley Kids series: adult ESL acquisition. *Journal of Reading*, 37, 662-67.

Chun, D. H., & Plass, J. L. (1996). Effects of multimedia annotations on vocabulary acquisition. *The Modern Language Journal*, 80, 183-198.

Coady, J. (1997). L2 vocabulary acquisition through extensive reading. *Second language vocabulary acquisition: A rationale for pedagogy*, 225.

Cohen, J. (Ed.). (1988). *Statistical Power Analysis for the Behavioral Sciences*. Lawrence Erlbaum Associates : New Jersey.

Day, R., Omura, C., & Hiramatsu, M. (1991). Incidental EFL vocabulary learning and reading. *Reading in a Foreign Language*, 7, 541–551.

Denscombe, M. (2007). *Educational research: Contemporary issues and practical approaches* (3rd ed.). London: Continuum.

Dubin, F., & Olshtain, E. (1993). Predicting word meanings from contextual clues: Evidence for L1 readers. In T. Huckin, M. Haynes, & J. Coady (Eds.), *Second language reading and vocabulary learning* (pp. 181–201). Norwood, NJ: Ablex.

Duncan, L. G., & Seymour, P. H. (2003). How do children read multisyllabic words? Some preliminary observations. *Journal of Research in Reading*, 26(2), 101-120.

Duncan, L. G., Seymour, P. H., & Bolik, F. (2007). Rimes and superrimes: an exploration of children's disyllabic rhyming skills. *British Journal of Psychology*, 98(2), 199-221.

Dupuy, B., & Krashen, S. (1993). Incidental vocabulary acquisition in French as a

foreign language. *Applied Language Learning*, 4(1&2), 55–63.

Elley, W. B. (1989). Vocabulary acquisition from listening to stories. *Reading Research Quarterly*, 24, 174–187

Elley, W. B. (1991). Acquiring literacy in a second language: the effect of book-based programs. *Language Learning* 41(3), 375-411.

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

Erdfelder, E., Faul, F., & Buchner, A. (1996). GPOWER: A general power analysis program. *Behavior research methods, instruments, & computers*, 28(1), 1-11.

Fraser, C. A. (1999). Lexical processing strategy use and vocabulary learning through reading. *Studies in Second Language Acquisition*, 21, 225–241.

Ghazal, Lotfi. (2010). Learning vocabulary in EFL contexts through vocabulary learning strategies. *Novitas-Royal*, 1(2), 84-91.

Godfroid, A., Boers, F., & Housen, A. (2013). An eye for words. *Studies in Second Language Acquisition*, 35(03), 483-517.

Gu, Y., & Johnson, R. (1996). Vocabulary learning strategies and language learning outcomes. *Language Learning*, 46, 643-679.

Hatch, E., & Brown, C. (1995). *Vocabulary, semantics, and language education*, Cambridge: Cambridge University Press.

Heatley, A., Nation, I.S.P. & Coxhead, A. (2002). RANGE and FREQUENCY programs. Available at <http://www.victoria.ac.nz/lals/staff/paul-nation.aspx>.

Holley, F. M., & King, J. K. (1971). Imitation and Correction in Foreign Language Learning<sup>1</sup>. *The modern language journal*, 55(8), 494-498.

Huckin, T., & Coady, J. (1999). Incidental vocabulary acquisition in a second language. *Studies in second language acquisition*, 21(02), 181-193.

Hudson, T. (2007). *Teaching second language reading*. New York, NY: Oxford University Press.

Hulstijn, J. H. (1992). Retention of inferred and given word meanings: Experiments in incidental vocabulary learning. In P. J. Arnaud & H. Bejoint (Eds.), *Vocabulary and applied linguistics* (pp. 113- 125). London: Macmillan.

Hulstijn, J. H. (1993). When do foreign-language readers look up the meaning of unfamiliar words? The influence of task and learner variables. *The Modern Language Journal*, 77(2), 139-147.

Hulstijn, J. H. (2003). Incidental and intentional learning. In C. J. Doughty & M. H. Long (Eds.), *Handbook of second language research* (pp. 349–381). Malden, MA: Blackwell.

Hulstijn, J., Hollander, M., & Greidanus, T. (1996). Incidental vocabulary learning by advanced foreign language students: The influence of marginal glosses, dictionary use, and reoccurrence of unknown words. *The Modern Language Journal*, 80(3), 327-339.

Horst, M., Cobb, T., & Meara, P. (1998). Beyond a clockwork orange: acquiring second language vocabulary through reading. *Reading in a Foreign Language*, 11: 207-23.

Jacobs, G., Dufon, P., & Hong, F.(1994). L1 and L2 vocabulary glosses in L2 reading passages: Their effects for increasing comprehension and vocabulary knowledge.



Journal of Research in Reading, 17(1), 19-28.

Kelch, K. (1985). Modified input as an aid to comprehension. *Studies in Second Language Acquisition*, 7, 81–90.

Kim, Youngkyu. (2006). Effects of input elaboration on vocabulary acquisition through reading by Korean learners of English as a foreign language. *TESOL Quarterly*, Vol. 40, No. 2.

Knight, S. (1994). Dictionary: The tool of last resort in foreign language reading? A new perspective. *The Modern Language Journal*, 78, 285-299.

Ko, M. H. (1995). Glossing in incidental and intentional learning of foreign language vocabulary and reading. *University of Hawaii Working Papers in ESL*, 13, 49–94.

Ko, M. H. (2005). Glosses, Comprehension, and Strategy Use. *Reading in a foreign language*, 17(2), 125-143.

Ko, M. H. (2012). Glossing and second language vocabulary learning. *Tesol Quarterly*, 46(1), 56-79.

Krashen, S. (1989). We acquire vocabulary and spelling by reading: Additional evidence for the input hypothesis. *The modern language journal*, 73(4), 440-464.

Kroll, J. F., & Curley, J. (1988). Lexical memory in novice bilinguals: The role of concepts in retrieving second language words. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds). *Practical Aspects of Memory: Current Research and Issues*, Vol. 2. (pp. 389-395). London: John Wiley & Sons.

Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory

representations. *Journal of Memory and Language*, 33(2), 149-74.

Kroll, J. F., & Sunderman, G. (2003). Cognitive processes in second language learners and bilinguals: The development of lexical and conceptual representations. *The handbook of second language acquisition*, 104-129.

Laufer, B. (1986). Possible changes in attitude towards vocabulary acquisition research. *IRAL-International Review of Applied Linguistics in Language Teaching*, (24), 69-75.

Laufer, B. (1990). Why are some words more difficult than others? Some intralexical factors that affect the learning of words. *International Review of Applied Linguistics in Language Teaching*, 28(4), 293-307.

Laufer, B., & Hill, M. (2000). What lexical information do L2 learners select in a CALL dictionary and how does it affect word retention? *Language Learning and Technology*, 3(2), 58-76.

Laufer, B., & Nation, I.S.P. (2012). Vocabulary. In A. Mackey, & S. M. Gass (Eds.), *The Routledge Handbook of Second Language Acquisition* (pp.163-176). Routledge.

Laufer, B., & Shumueli, K. (1997). Memorizing new words: Does teaching have anything to do with it? *RELC Journal*, 28(1), 89-108.

Lyman-Hager, M., Davis, J. N., Burnett, J., & Chennault, R. (1993). Une vie de boy: Interactive reading in French. In F. L. Borchardt & E. M. T. Johnson (Eds.), *Proceedings of the CALICO 1993 Annual Symposium on Assessment* (pp. 93-97). Durham, NC: Duke University.

McCarthy, M., & Carter, R. (1997). Grammar, tails, and affect: Constructing expressive choices in discourse. *Text-Interdisciplinary Journal for the Study of Discourse*, 17(3), 405-429.

Meara, P. (1984). The study of lexis in interlanguage. In A. Davies, C. Criper, & A. P. R. Howatt (Eds.), *Interlanguage* (pp. 225-239). Edinburgh, Scotland: Edinburgh University Press.

Milton, J., & Treffers-Daller, J. (2013). Vocabulary size revisited: the link between vocabulary size and academic achievement. *Applied Linguistics Review*, 4(1), 151-172.

Miyasako, N. (2002). Does text-glossing have any effects on incidental vocabulary learning through reading for Japanese senior high school students?. *Language Education & Technology*, (39), 1-20.

Nagata, N. (1999). The effectiveness of computer-assisted interactive glosses. *Foreign Language Annals*, 32, 469-479.

Nagy, W. (1997). On the role of context in first- and second-language vocabulary learning. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary description, acquisition and pedagogy* (pp. 64–83). New York, NY: Cambridge University Press.

Nagy, W. E., & Anderson, R. C. (1984). How many words are there in printed school English?. *Reading Research Quarterly*, 304-330.

Nagy, W. E., Anderson, R. C., & Herman, P. A. (1987). Learning word meanings from context during normal reading. *American educational research journal*, 24(2), 237-270.

Nagy, W. E., & Herman, P. (1985). Incidental vs. Instructional approaches to increasing reading vocabulary. *Educational Perspectives*, 23, 16-21.

Nagy, W. E., & Herman, P. (1987). Breadth and depth of vocabulary knowledge: Implications for acquisition and instruction. In M. McKeown & M. Curtis (Eds.), *The nature of vocabulary acquisition* (pp. 19-35). Hillsdale, NJ: Lawrence Erlbaum.

Nagy, W. E., Herman, P., & Anderson, R. (1985). Learning words from context. *Reading Research Quarterly*, 20, 233-253.

Nation, I.S.P. (1990). *Teaching and Learning Vocabulary*. New York: Newbury House.

Nation, I. S. P. (2001). *Learning vocabulary in another language*. New York, NY: Cambridge University Press.

Nation, I.S.P. (2010). *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press.

Nation, P., & Coady, J. (1988). Vocabulary and reading. *Vocabulary and language teaching*, 97, 110.

Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational psychologist*, 38(1), 1-4.

Paribakht, T., & Wesche, M. (1996). Enhancing vocabulary acquisition through reading: A hierarchy of text-related exercise types. *The Canadian Modern Language Review*, 52(2), 155-178.

Paribakht, T., & Wesche, M. (1997). Vocabulary enhancement activities and reading for meaning in second language vocabulary acquisition. In J. Coady, & T. Huckin, (Eds.) *Second Language Vocabulary Acquisition* (pp. 174-200). Cambridge: Cambridge University Press.

Park, E. S. (2005). Constraints of implicit focus on form: Insights from a study of input enhancement. *Teachers College, Columbia University Working Papers in TESOL & Applied Linguistics*, 4(2).

Parry, K. (1993). Too many words: Learning the vocabulary of an academic subject. In T. Huckin, M. Haynes, & J. Coady (Eds.), *Second language reading and vocabulary learning* (pp. 109–129). Norwood, NJ: Ablex.

Parry, K. (1997). Vocabulary and comprehension: Two portraits. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition: A rationale for pedagogy* (pp. 55-68). New York: Cambridge University Press.

Pignot-Shahov, V. (2012). Measuring L2 receptive and productive vocabulary knowledge. *Language Studies Working Papers*, 4, 37-45.

Pitt, M., White, H., & Krashen, S. (1989). Acquiring second language vocabulary through reading. *Reading in a foreign language*, 5, 271-275.

Potter, M. C., So, K., Eckardt, V., & Feldman, L. B. (1984). Lexical and conceptual representation in beginning and proficient bilinguals. *Journal of Verbal Learning and Verbal Behavior*, 23(1), 23-38.

Rastle, K., Harrington, J., & Coltheart, M. (2002). 358,534 nonwords: The ARC nonword database. *The Quarterly Journal of Experimental Psychology: Section A*, 55(4), 1339-1362.

Rott, S., Williams, J., & Cameron, R. (2002). The effect of multiple-choice L1 glosses and input-output cycles on lexical acquisition and retention. *Language Teaching Research*, 6, 183-222.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11, 129–158.

Schmidt, R. (1993). Awareness and second language acquisition. *Annual Review of Applied Linguistics*, 13, 206-226.

Schmidt, R. (1994). Deconstructing consciousness in search of useful definitions for applied linguistics. *AILA Review*, 11, 11-26.

Schmidt, R. W. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction*. Cambridge: Cambridge University Press, 3-32.

Schmitt, N. (2013). *An introduction to applied linguistics*. Routledge.

Stein, J. (1997). How institutions learn: a socio-cognitive perspective. *Journal of Economic Issues*, 729-740.

Swanborn, M. S. L. & De Glopper, K. (1999). Incidental word learning while reading: A meta-analysis. *Review of Education Research*, 69, 261-85.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive science*, 12(2), 257-285.

Taylor, A. M. (2010). Glossing Is Sometimes a Distraction: Comments on Cheng and Good (2009). *Reading in a Foreign Language*, 22(2), 353-354.

Thorndike, E. L. (1914). *Educational psychology*, Vol 3: Mental work and fatigue and individual differences and their causes.

Watanabe, Y. (1997). Input, intake, and retention: Effects of increased processing on incidental learning of foreign language vocabulary. *Studies in Second Language Acquisition*, 19(3), 287-307.

Wode, H. (1999). Incidental vocabulary acquisition in the foreign language classroom. *Studies in Second Language Acquisition*, 21, 243–258.

Wong, W. (2005). *Input enhancement: From theory and research to the classroom*. New York, NY: McGraw-Hill.

Yeh, Y., & Wang, C. (2003). Effects of multimedia vocabulary annotations and learning styles on vocabulary learning. *CALICO Journal* 21(1), 131-144.

Yoshii, M. (2006). L1 and L2 glosses: Their effects on incidental vocabulary learning. *Language Learning & Technology*, 10(3), 85-101.

Yoshii, M., & Flaitz, J. (2002). Second language incidental vocabulary retention: The effect of text and picture annotation types. *CALICO journal*, 20(1), 33-58.

Yoshii, M. (2013). Effects of gloss types on vocabulary learning through reading: Comparison of single and multiple gloss types. *CALICO Journal*, 30, 203-229.

Zimmerman, C. B. (1997). Do reading and interactive vocabulary instruction make a difference? An empirical study. *TESOL quarterly*, 31(1), 121-140.