THE EFFECT OF FIRST-LANGUAGE VERSUS SECOND-LANGUAGE GLOSSES ON INCIDENTAL VOCABULARY LEARNING AND RETENTION

İREM ORHAN KONUK

BOĞAZİÇİ UNIVERSITY

2019

THE EFFECT OF FIRST-LANGUAGE VERSUS SECOND-LANGUAGE GLOSSES ON INCIDENTAL VOCABULARY LEARNING AND RETENTION

Thesis submitted to the

Institute for Graduate Studies in Social Sciences

in partial fulfillment of the requirements for the degree of

Master of Arts

in

English Language Education

by

İrem Orhan Konuk

Boğaziçi University

2019

The Effect of First-Language Versus Second-Language Glosses

on Incidental Vocabulary Learning and Retention

The thesis of İrem Orhan Konuk

has been approved by:

Prof. Gülcan Erçetin
(Thesis Advisor)

Assist. Prof. Sibel Tatar

Assist. Prof. Mustafa Polat
(External Member)

August 2019

DECLARATION OF ORIGINALITY

I, İrem Orhan Konuk , certify that

- I am the sole author of this thesis and that I have fully acknowledged and documented in my thesis all sources of ideas and words, including digital resources, which have been produced or published by another person or institution;
- this thesis contains no material that has been submitted or accepted for a degree or diploma in any other educational institution;
- this is a true copy of the thesis approved by my advisor and thesis committee at Boğaziçi University, including final revisions required by them.

Signature İrem Konuk

Date 06.08.2019

ABSTRACT

The Effect of First-Language Versus Second-Language

Glosses on Incidental Vocabulary Learning and Retention

The study investigates the effects of different type of computer-mediated textual glosses and no-gloss condition on incidental vocabulary learning and retention. Seventy-four native speakers of Turkish studying English at a Turkish state university participated the study. Participants were randomly assigned to three conditions: (i) Turkish (L1) gloss, (ii) English (L2) gloss, and (iii) no gloss. First, they were asked to read an English text for general comprehension and to use the glosses to enhance reading comprehension. The data were collected using both quantitative and qualitative measures. Upon reading the text, participants were administered a multiple-choice reading comprehension test as it was announced at the beginning. Next, they were administered an unannounced (in order to measure incidental vocabulary learning) on-screen vocabulary post-test. Lastly, subjects were asked to answer a questionnaire. Three weeks later they were administered an unannounced delayed vocabulary test, and volunteers were interviewed. Results of the study indicate that gloss use did not facilitate reading comprehension. Both L1 and L2-glosses were found to be effective in incidental vocabulary learning, L1-glosses being more effective than L2-glosses. Findings of the study also suggest that gloss condition does not affect vocabulary retention. Tracking data indicated that incidental vocabulary learning is positively related to the frequency of access to glosses as well as with the time spent accessing glosses. According to the qualitative data, majority of the participants preferred having L2-glosses in their reading texts and they perceived glosses to be useful and timesaving.

ÖZET

Anadilde ve Yabancı Dilde Hazırlanmış Açıklama Notlarının Tesadüfi Kelime Öğrenimine ve Kelimelerin Akılda Tutulabilmesine Etkileri

Bu çalışma anadilde (L1) ve yabancı dilde (L2) hazırlanmış açıklama notlarının okuma yoluyla tesadüfi kelime öğrenimine ve kelimelerin akılda tutulmasına etkisini bilgisayar ortamında incelemiştir. Çalışmaya anadili Türkçe olan ve Türkiye'deki bir devlet üniversitesinde yabancı dil olarak İngilizce eğitimi almakta olan 74 öğrenci katılmıştır. Öğrenciler rastgele atama yolu ile üç gruba ayrılmıştır: (i) Türkçe (L1) açıklama, (ii) İngilizce (L2) açıklama, ve (iii) açıklamasız. İlk olarak öğrencilerden İngilizce bir metni kavrama amaçlı olarak okumaları ve açıklama notlarını da kavramalarına yardımcı olması için kullanmaları istenmiştir. Araştırmada hem nicel hem de nitel veri toplanmıştır. Metin okunduktan sonra öğrencilere çoktan seçmeli bir kavrama testi uygulanmıştır. Daha sonra bir kelime bilgisi son testi habersiz olarak uygulanmıştır. Son olarak öğrencilerden bir anket doldurmaları istenmiştir. Üç hafta sonra öğrencilere yine habersiz bir gecikmeli son test uygulanmış arkasından da gönüllü öğrencilerle mülakat yapılmıştır. Sonuçlar, açıklama notu kullanımının okuduğunu kavrama üzerinde kolaylaştırıcı bir etkisi olmadığını göstermiştir. Hem anadildeki hem de yabancı dildeki açıklama notlarının tesadüfi kelime öğreniminde etkili oldukları, anadildeki açıklamaların yabancı dildeki açıklamalara kıyasla daha etkili olduğu anlaşılmıştır. Araştırmanın sonuçlarına göre açıklama notu kullanmanın kelimelerin akılda tutulmasına etkisi bulunmamaktadır. Tesadüfi kelime öğreniminin, notlara erişme sıklığı ve notların inceleme süresi ile pozitif yönlü bir ilişkisinin bulunduğu anlaşılmıştır. Katılımcıların çoğunluğu okuma metinlerinde yabancı dilde hazırlanmış açıklama notlarını kullanmayı tercih edeceklerini ifade etmiştir.

ACKNOWLEDGMENTS

I would first like to thank my thesis advisor Prof. Gülcan Erçetin of the Foreign Language Education Department at Boğaziçi University. The contribution of her expertise, profound knowledge and experience to my study was invaluable. I am thankful to her for all that she did.

I am grateful to Assist. Prof. Elif Kemaloğlu Er for making it possible for me to collect data by welcoming me to her classes. I am thankful to her for her positive attitude, kindness and friendliness.

I would like also to thank my mother Fatma, my father Mehmet, and my brother Kerem for loving me unconditionally and being there for me whenever I turn to. I am also thankful to my dear friend Eda Üzüm for believing in me and motivating me at moments of despair.

Finally, I would like to thank my dear husband Melih for supporting me in my decisions and choices no matter what the circumstances are. I am also thankful to my son Mehmet Erdem and my daughter Bilge for bringing so much joy and love to my life and teaching me the things that I thought I knew anew.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION 1

1.1 Purpose of the study 5

1.2 Significance of the study 6

CHAPTER 2: LITERATURE REVIEW 7

2.1 Gloss use and reading comprehension 7

2.2 Effect of L1 versus L2-glosses on incidental vocabulary learning 12

2.3 Effect of L1 versus L2-glosses on vocabulary retention 16

2.4 The present study 20

CHAPTER 3: METHODOLOGY 21

3.1 Participants 21

3.2 Materials and instruments 21

3.3 Procedures 27

3.4 Data analysis 28

CHAPTER 4: RESULTS 30

4.1 Effects of glosses on reading comprehension 30

4.2 Incidental vocabulary learning 30

4.3 Analyses based on log data 35

4.4 Insights from the questionnaires and interviews 37

CHAPTER 5: DISCUSSION AND CONCLUSION 42

5.1 Pedagogical implications 47

5.2 Limitations and suggestions for further research 48

APPENDIX A: READING TEXT 49

APPENDIX B: GLOSSED WORDS 51

APPENDIX C: COMPREHENSION TEST 52

APPENDIX D: FORM RECOGNITION TEST 54

APPENDIX E: MEANING PRODUCTION TEST 55

APPENDIX F: MEANING RECOGNITION TEST 56

APPENDIX G: STUDENT OPINION QUESTIONNAIRE 58

APPENDIX H: INTERVIEW QUESTIONS 59

REFERENCES 60

LIST OF FIGURES

Figure 1. A screenshot showing how glosses are presented in the software 22

Figure 2. A screenshot showing the screen layout of the reading text screen 23

Figure 3. Data collection procedures of the present study 28

Figure 4. Interaction between vocabulary knowledge and time 33

Figure 5. Interaction between gloss condition and time 34

ABBREVIATIONS

CALL: Computer Assisted Language Learning

EFL: English as a Foreign Language

ESL: English as a Second Language

L1: First-Language

L2: Second-Language

CHAPTER 1

INTRODUCTION

The incorporation of authentic texts into the second language curriculum brought with itself an overburden of unknown vocabulary on the part of the lower-level language learners. For instance, Lyman-Hager and Davis (1996) stated that intermediate-level French language students experienced difficulties and reported dislike after reading authentic passages from French literature. The most prominent cause for this was the large number of unfamiliar words that the students encountered while reading (Lyman-Hager & Davis, 1996). There are a number of strategies that students may choose from to overcome this problem of unknown vocabulary such as: using dictionaries or trying to guess the meaning of the word from its context. According to Nation and Meara (2010) inferring a meaning for a word from contextual cues is the most useful strategy of all. Yet, in order to be able to use this strategy effectively, learners need to know the meaning of 95–98 per cent of the words in a text (Nation & Meara, 2010). It is clear that intermediate-level second-language (L2) learners would probably not benefit from this strategy due to their limited vocabulary size. Therefore, in order to make authentic reading materials more accessible to lower-level language learners, input modification techniques may be more helpful.

Use of glosses is a common type of input modification that helps to make authentic texts accessible to language learners (Bell & LeBlanc, 2000). Gloss is a short definition that is provided to help L2 learners with their reading and comprehension processes (Lomicka, 1998). They are provided generally for unfamiliar words (Lomicka, 1998) in order to ensure an uninterrupted reading

process (Rott & William, 2003). By manipulating the target form, glossing helps learners, to notice and learn the meaning of unknown words when contextual cues are not sufficient for them to infer the meaning by themselves (Ko, 2012). Though necessary, exposure to input alone is not sufficient for learners to form solid lexical knowledge which can later be put into use actively and appropriately. Schmidt's (1990, 1993) noticing theory, places attention to a crucial position in the language learning process. In the same vein, Nation and Meara (2010) hold that glossing and highlighting the unfamiliar vocabulary items would increase learners' deliberate attention to those words, which in turn would increase vocabulary learning. Hulstijn (2001) also refers to a number of empirical studies suggesting that acquisition of L2 vocabulary while reading will be considerably increased if learners' attention is directed towards unfamiliar words, by means of marginal glosses or with tasks demanding words to be looked up in a dictionary, or with simple fill-in exercise. The fact that glosses are boldfaced also prevent learners from ignoring target lexical items as part of an avoidance strategy that is used when a word is not deemed crucial for general comprehension (Schmitt, 1997).

A number of Computer Assisted Language Learning (CALL) reading studies investigated the effectiveness of providing immediate access to the meanings of unfamiliar vocabulary by means of online glossing. According to Chun (2006), online glossing allows language learners reach the meaning of unfamiliar words fast and easily. Online glossing compensates for learners' insufficient automatic lower-level processes, and by doing so, it enables learners to pay attention to higher level processes (Chun, 2006). Indeed, online glosses are easily accessed by clicking on the unfamiliar word, which facilitates interactivity and helps learners to be actively involved in the reading process.

The unobtrusive nature of online glosses incorporated into computer mediated texts do not only facilitate the reading process for the learner by making reading more pleasurable and fluent (Lyman-Hager & Davis, 1996) but also help researchers understand the reading process by tracking the choices made by individual readers during reading. Thus, researchers can unobtrusively attain in depth information about both the product and the process of learning (Collentine, 2000), which would not be possible with marginal glosses incorporated into printed texts.

In paper-based studies, researchers have to rely on learners' reports as to their use of glosses, the amount of time spent on using glosses or reading the text, making the task become labor intensive (Davis, 1995) and lead to imprecise and unreliable results.

Incidental vocabulary learning, which refers to acquisition of vocabulary as the byproduct of any activity not intentionally aimed at vocabulary learning (Hulstijn, 2001) has been studied by a number of studies in relation to the use of glosses during reading and listening tasks. In studies investigating incidental vocabulary learning, learners are generally requested to read a text involving processing of some unfamiliar words without being announced that they would subsequently be tested on their recall of those words. Whereas, in studies investigating intentional learning, learners are told in advance that they will be administered a test afterwards (Laufer & Hulstijn, 2001).

Miyasako (2002) asserts that the rationale behind incidental vocabulary learning can be explained on the basis of Input Hypothesis. According to Input Hypothesis unfamiliar lexical items can automatically be acquired if learners attend to the meaning of comprehensible input (Pitts, White & Krashen, 1989). The findings of Pitts et al. (1989) revealed that it is possible to attain an increase in one's lexical

knowledge by reading for meaning. Therefore, second language learners are recommended to engage in extensive reading activities. Yet, relying only on reading for vocabulary growth is not an efficient strategy for language learners. Schmidt (2001) argued that, if learners take a passive approach to learning and depend solely on unintentional processes to bring about automatic noticing, they will probably not be successful (as cited in Laufer & Hulstijn, 2001, p. 4). Here glosses seem to be a useful option to help learners notice the target lexical items while reading.

The type of lexical information provided by glosses such as first-language (L1) translation or L2 definition has been of particular concern in the L2 studies. This concern originates from the language decisions made on whether or not to use L1 in the foreign or second language classrooms. In English as a Foreign Language (EFL) environments learners have little if any opportunities to get exposed to the target language other than what is provided in their language lessons. Therefore, teachers aim at maximizing the use of L2 in the classroom including the use of monolingual dictionaries. However, in order to effectively use such dictionaries, learners need to have a target language vocabulary of at least 2,000 words (Nation, 2003). Nation (2003) claims that most EFL learners manage to acquire this much vocabulary only after five to six years of studying English.

According to the literature there are certain conditions in which use of either L1 translations or L2 definitions for learning the meaning of an unknown word might be more advantageous for the learner than the other. For instance, the revised hierarchical model proposed by Kroll and Stewart (1994) assumes that in earlier stages of L2 acquisition, L2 words are associated with their L1 translations in order to access the meaning that is already represented for those concepts. As L2 proficiency increases, direct conceptual connections from L2 words to semantics

begins to develop (Kroll & Sunderman, 2003). Hulstijn (2001) also asserts that low-level L2 learners mostly seem to associate the L2 word form directly with a comparable L1 word form while they acquire their first few hundred target language vocabulary items. "Thus, initially, L2 lexical entries are often coded as phonological or orthographic extensions of L1 lexical entries" (Hulstijn, 2001, p. 4). L2 learners associate the L2 word form directly with its meaning only when they reach a later stage in proficiency (Hulstijn, 2001). In the same vein Taylor (2006) states that learners whose proficiency level is intermediate and above feel comfortable consulting L2-glosses.

1.1 Purpose of the study

There have been inconsistent findings in previous second language research on the effect of textual glosses on incidental vocabulary learning (Jacobs et al., 1994; Chen, 2002; Yoshii, 2006; Ko, 2012; Miyasako, 2002; Arpacı, 2016), retention (Laufer & Shmueli, 1997; Arpacı, 2016; Jacobs et al., 1994; Miyasako, 2002; Yoshii, 2006; Ko, 2012), and reading comprehension (Johnson, 1982; Pak, 1986; Jacobs, Dufon, & Hong, 1994; Davis & Lyman-Hager, 1997; Miyasako, 2002; Ariew & Ercetin, 2004; Akbulut, 2007; Al-Jabri, 2009; Cheng & Good, 2009; Davis, 1989; Jacobs, 1994; Ko, 2005; Azari, 2012; Hashemian & Fadaei, 2013; Arpacı, 2016). The contradictory results call for additional studies to extend this body of research by including different groups of second language learners, different reading texts, and different groups of glossed vocabulary items, which would shed light on the probable causes of the contradictory results and would hopefully lead to more generalizable findings. The current study aims at extending this body of research as mentioned above and by

providing additional types of data (time spent on using the glosses and frequency of access to glosses) thanks to the tracking option of the reading software that was used.

1.2 Significance of the study

The results derived from the study will provide insights which will help material designers, teachers and learners to make learned decisions in terms of textual gloss use. For instance, the quantitative data on; the time spent on using the glosses and the frequency of access to glosses allow for making analyses to answer questions such as: (i) Does the type of gloss affect the number of glosses accessed and the amount of time spent on using the glosses? (ii) Is the frequency of access to glosses and the amount of time spent on glosses related to incidental vocabulary learning and reading comprehension? Answers to these questions may help teachers and learners in their choices of whether to use glosses, which type of glosses to use, and how to use them.

The qualitative data gathered by means of a questionnaire and interviews will depict upper-intermediate level EFL learners' preferences and perceptions about gloss use, gloss language, vocabulary learning, and reading from a computer screen. These data along with the results of the analyses of quantitative data will help discover and eliminate any misconception on the minds of the learners pertaining to afore mentioned issues. More individualized reading software may be developed by taking these learner preferences into consideration. By doing so individual learner differences would also be accounted for.

CHAPTER 2

LITERATURE REVIEW

2.1 Gloss use and reading comprehension

Many foreign language educators and students tend to believe that glosses facilitate reading comprehension. However, research on the effectiveness of glosses produced inconsistent results. Findings of some studies (Johnson, 1982; Pak, 1986; Jacobs, Dufon, & Hong, 1994; Davis & Lyman-Hager, 1997; Miyasako, 2002; Ariew & Ercetin, 2004; Akbulut, 2007; Al-Jabri, 2009; Cheng & Good, 2009) suggested no significant effect for glossing in L2 reading comprehension, whereas some other studies (Davis, 1989; Jacobs, 1994; Ko, 2005; Azari, 2012; Hashemian & Fadaei, 2013; Arpacı, 2016) came up with research results implying that glosses do facilitate reading comprehension.

Davis and Lyman-Hager (1997) investigated the relationship between reading comprehension and use of computerized glosses. Participants were 42 intermediate-level undergraduate students of French. Participants read the text from the computer screen and they had access to five types of glosses. The gloss types included: word definitions in English and in French, pronunciation, grammar, referent explanation, cultural explanations involving both pictures and texts. Reading comprehension was operationalized as performance on a multiple-choice test and a recall protocol in English. The results of the study showed that the participants accessed English definitions more than any other gloss types. However, the results suggested that textual gloss use does not enhance reading comprehension. This was contradictory with the participants' perceptions. Participants felt that gloss use helped them

comprehend better, but their test scores did not yield any evidence that gloss use enhanced their comprehension.

Ariew and Ercetin's (2004) study investigated the effects of picture, video, and sound annotations as well as textual glosses on reading comprehension. 84 intermediate and advanced level adult intermediate and advanced-level English as a Second Language (ESL) learners participated the study. The participants' reading comprehension was measured through a 14-item test involving six short-answer questions, six multiple choice questions, and two open-ended questions. The results suggested that textual gloss use did not facilitate reading comprehension for either group.

Akbulut's (2007) research results also confirm Davis and Lyman-Hager's (1997) and Ariew and Ercetin's (2004) findings. Akbulut's (2007) participants were 69 freshman students at a Turkish university. They were advanced-level learners of English. Participants were randomly assigned to one of three gloss conditions: (i) word definitions in English, (ii) definitions along with related pictures, and (iii) definitions along with short videos. Upon reading an annotated text they were administered a reading comprehension test with two sections: multiple-choice and true-false. The results suggested no relationship between gloss use and reading comprehension.

There have also been a number of studies which compared the effectiveness of L1 versus L2-glosses on L2 reading comprehension (Jacobs et al., 1994; Bell & LeBlanc, 2000; Ko, 2005; Hashemian & Fadaei, 2013; Arpacı, 2016). For instance, Jacobs et al. (1994) investigated the effect of gloss use on recall and vocabulary learning, and learners' preferences as to glossing. Participants were 85 native speakers of English enrolled in fourth semester Spanish at Florida State University.

They read a Spanish text under three conditions: English (L1) gloss, Spanish (L2) gloss, and no gloss. Reading comprehension was operationalized as performance on a recall test. The test required participants to write in English as much of the text as they could recall. The results suggested no significant differences in reading comprehension scores resulted from the various gloss conditions.

Bell and LeBlanc (2000) investigated whether glosses in L1 or L2 are more frequently used by language learners during computer-based reading. Participants were 40 undergraduates enrolled in third semester Spanish classes. They read a Spanish text under two conditions: English (L1) gloss and Spanish (L2) gloss. Participants' reading comprehension was measured through a 10-item multiple choice test. The findings suggested no significant differences between the L1 and L2 groups in terms of their reading comprehension. Since the research design did not include a control group, the results do not provide information as to whether gloss use enhance reading comprehension or not.

Ko (2005) investigated the effectiveness of the language of glosses on reading comprehension. Participants were 106 undergraduates who were enrolled in the second semester of a freshman English class (they had six years of English instruction before entering university) at a university in Korea. There were three treatment conditions: Korean (L1) gloss, English (L2) gloss, and no gloss. Reading comprehension was measured through a multiple-choice test consisting of 25 items in English. Results revealed that only L2-glosses significantly affected students' reading comprehension. Ko (2005) asserted three possible explanations for the contradictory results between previous research and her study. First, she suggested that when deciding on which words to gloss, one should consider the relative importance of the words in understanding the text. Second, she asserted that for more

proficient learners who can comprehend the definitions in L2, L2-glosses might be more effective. However, Ariew and Ercetin's (2004) results suggested that textual (L2) gloss use did not facilitate reading comprehension for either intermediate or advanced ESL learners. Similarly, Akbulut (2007) also had advanced-level participants and the findings suggested no relationship between gloss use and reading comprehension. Another issue to point out is that the proficiency levels of the participants in Ko's (2005) study are unknown. The only information related to participants' proficiency level is that they were in their second semester in English, and that they had had six years of English instruction before entering university. Lastly, Ko (2005) stated that the way in which reading comprehension is measured can influence reading comprehension scores. She suggests that in order to get more reliable results there should be consistency in terms of how reading comprehension gets assessed.

Taylor (2006) suggested two additional explanations for Ko's (2005) finding. First, he stated that the language of the testing instrument might have an influence on the results. In the case of Ko (2005) both the text and the reading comprehension tests were written in the L2 (English). Therefore, Taylor (2006) concludes that it is quite reasonable for L2-gloss group to perform better. Secondly, he contends that if the text level is relatively lower than the participants' proficiency level, the participants might not have felt the need to consult L1-glosses at all. However, this second explanation does not really explain why L2-glosses were found to be effective in reading comprehension.

Arpacı (2016) conducted a study with 81 10^{th} graders studying at a high school in Turkey. Their proficiency level was lower-intermediate. The aim of the study was investigating the effectiveness of the language of glosses on second

language vocabulary learning. Participants read an English text under three conditions: Turkish (L1) gloss, English (L2) gloss, and no gloss. Reading comprehension was measured through a multiple-choice test consisting of 10 items in English. The results of the study suggested no significant difference between L1 and L2-gloss conditions in terms of reading comprehension, however a significant difference was observed between both gloss conditions and no gloss condition. This finding confirms the results of Jacobs' (1994) study conducted with 116 undergraduates enrolled in third semester Spanish. They were requested to read a Spanish text in two conditions: L1 (English) gloss and no gloss. Reading comprehension was operationalized as performance on a recall test. The results suggested that glossed group performed significantly better than the no gloss group. Jacobs (1994) stated that glossing should not be viewed as the sole factor which can affect comprehension. He contends that the effectiveness of glosses on reading comprehension might be determined by a number of different factors such as: the difficulty level of the text, the proficiency level of the readers, and the items that are chosen to be glossed. He also states that the reason why glossing appeared to be effective in his study might be the utilization of higher than average number of glossed items (glossed words constituted 11% of the text, however for Arpacı (2016) who reached similar results had a ratio of 6%).

To sum up, it can be claimed that the reason behind the mixed results in the literature seems to stem from the lack of comparability between the studies, in terms of; proficiency levels of the participants, difficulty levels of the texts, types of comprehension tests, and number of glossed items.

2.2 Effect of L1 versus L2-glosses on incidental vocabulary learning

The studies investigating the effectiveness of gloss language on L2 incidental vocabulary learning brought inconsistent results. A number of studies reported that no significant differences were found between L1 versus L2-glosses in terms of their effectiveness in L2 incidental vocabulary learning (Jacobs et al., 1994; Chen, 2002; Yoshii, 2006; Ko, 2012). There are also other studies (Miyasako, 2002; Arpacı, 2016) reporting significant differences between L1 versus L2-glosses in terms of their effectiveness in L2 incidental vocabulary learning.

Jacobs et al. (1994) investigated the effectiveness of L1 versus L2-glosses for increasing comprehension and vocabulary knowledge. Participants were asked to read a Spanish text under three conditions: English (L1) gloss, Spanish (L2) gloss, and no gloss. Upon reading the text, they were administered a vocabulary test involving translation of 32 glossed words from Spanish to English. Results of the study showed that the two gloss groups performed significantly better than the control group. However, no significant differences were found between scores of the English gloss group and Spanish gloss group. However, the testing instrument might have an influence on the results. Vocabulary acquisition is an incremental process (Herman, Anderson, Pearson & Nagy, 1987); learners may grasp only some portion of the meaning of an unknown vocabulary item at first encounter. Therefore, tests assessing incidental vocabulary learning should be sensitive to partial gains in word knowledge (Herman et al., 1987). A learner might not be able to translate a word after being exposed to it just once, but a multiple-choice test or a matching task may give her a clue and help her bring out the partial gain she got. This is to say, using a multiple-choice test or a matching task along with the translation task may prevent

researchers to conclude erroneously that incidental acquisition of vocabulary knowledge has not occurred.

Yoshii (2006) investigated the effectiveness of gloss language on incidental vocabulary learning in a multimedia environment. Participants consisted of 130 freshmen, 29 sophomores, and 36 juniors who were low-intermediate EFL learners in Japan. Participants were administered a pretest one week prior to the treatment. They were required to write down the L1 explanations of the words that they knew from among the given 20 words (14 target words and 10 distractors). One-week later participants were asked to read a computer-mediated English text in four conditions: (1) L1 text only; (2) L2 text only; (3) L1 text plus picture; and (4) L2 text plus picture. Upon reading the text participants received a vocabulary test consisting of two sections: a definition-supply and a recognition test. The format of the definition-supply test was the same with the pretest. Recognition test involved 14 multiple-choice questions requiring participants to select the most appropriate definition of a given word from four choices written in L2 (English). Although results of the study suggested that both L1 and L2-glosses are effective for incidental vocabulary learning, neither the definition-supply nor the recognition tests revealed significant differences between L1 and L2 text-only gloss types.

Ko (2012) investigated the effectiveness of different types of glosses and no gloss condition on L2 vocabulary learning. Participants were 90 freshmen students at a university in South Korea. They were intermediate-level learners of English. Participants were requested to read an English text under three different gloss conditions: Korean (L1) gloss, English (L2) gloss, and no gloss. Upon reading the text they were administered an unannounced multiple-choice vocabulary test. Results

suggested that the glossed groups performed significantly better than no gloss group. However, post-test scores of L1 and L2-gloss groups did not differ significantly.

Miyasako (2002) compared the effectiveness of L1 versus L2-glosses and also investigated the effectiveness of multiple-choice versus single glosses. Participants were 187 third year students at a senior high school in Japan. They were judged to be pre-intermediate or intermediate level learners of English. Study involved six conditions: (i) L2 (English) multiple-choice gloss, (ii) L1 (Japanese) multiple-choice gloss; (iii) L2 single gloss; (iv) L1 single gloss; (v) No gloss; and (vi) Control (no reading). Upon reading an English text, participants took a 15-item multiple-choice vocabulary test. The results revealed that the L2-gloss groups performed significantly better than the L1-gloss groups on the immediate test. The results also suggested that L2-glosses were more effective for higher proficiency level learners, whereas L1-glosses were more effective for learners with lower-proficiency levels.

Cheng & Good (2009) investigated the effects of different types of glosses on reading comprehension and vocabulary acquisition. Participants were 135 undergraduate business and engineering students at four English proficiency levels studying at a technical university in Taiwan. The proficiency levels were designated as Levels 1 to 4 from the lowest to the highest level on the basis of scores taken from English Placement Test developed by the University of Michigan. After completing a pretest (they matched L2 target words with their L1 translations), participants were asked to read an English text under four conditions: (i) L1-gloss-L2-example, (ii) L1-in-text-gloss, (iii) L1-MG-gloss, (iv) no-gloss. Upon reading, the participants were requested to complete a fill in the blanks type vocabulary test. The test involved 16 English sentences each with a blank that students were required to fill in choosing

the most appropriate word from a list of 24 English items. The results of the study suggested a significant difference between glossed groups and no gloss group. Although there is no mention of a significant difference between the vocabulary test scores of the different gloss conditions, L1-gloss-L2-example gloss group scored the highest followed by L1-in-text-gloss group and L1-MG-gloss group scored the lowest.

Arpacı (2016) investigated the effectiveness of L1 (Turkish) versus L2 (English) glosses and no gloss condition on incidental vocabulary learning. 81 lower-intermediate level 10th graders studying at a high school in Turkey were asked to read an English text under three conditions: Turkish (L1) gloss, English (L2) gloss, and no gloss. Upon reading the text an unannounced vocabulary test was administered. The test consisted of 16 multiple-choice questions. The results of the study suggested a significant difference between glossed groups and no gloss group. The vocabulary test results of L1 and L2-gloss groups also differed significantly, favoring L1-gloss condition.

There seems to be a consensus in the literature on the effectiveness of gloss conditions over no gloss condition on incidental vocabulary learning. However, it is not clear whether language of gloss also has an effect on incidental vocabulary learning. According to Miyasako (2002) and Arpacı (2016) there is a significant difference between L1 and L2-gloss groups. In Miyasako (2002) the difference favors L2-gloss group, whereas in Arpacı (2016) it is L1-gloss group who performed significantly better. However, Miyasako's (2002) results also revealed that L1-glosses were more effective for lower-proficiency level learners. The participants of her study consisted of pre-intermediate or intermediate level learners. Participants of Arpacı (2016), on the other hand, were lower-intermediate level EFL learners. By

looking at this information, we may conclude that the findings of Arpacı (2016) confirms Miyasako's (2002) results suggesting L1-glosses were more effective for lower-proficiency level learners. However, the proficiency level of the participants in both studies were not determined on the basis of a standardized testing instrument. Miyasako (2002) stated that "Their English proficiency was judged to be in a pre-intermediate or intermediate level after five years of EFL education" (p. 5). Arpacı (2016), on the other hand, accepted the proficiency level reported by the participants' school administration which is reported to be A2 based on Common European Framework. This raises a question mark on the comparability of the two research results on the basis of proficiency levels.

2.3 Effect of L1 versus L2-glosses on vocabulary retention

The research results reviewed here all reported that there was a decrease in the delayed post-test scores of their participants. Which is not surprising when we consider that vocabulary acquisition is a cumulative (Nation & Meara, 2010) and an incremental (Herman, Anderson, Pearson & Nagy, 1987) process. According to Nation & Meara (2010) the knowledge of the word becomes stronger only after getting exposed to the target vocabulary items a number of times. Although it is not clear as to how many times a word needs to be encountered for incidental vocabulary learning to occur through reading, Wodinsky and Nation (1988) suggests 10 repetitions. Nagata (1999) points out the importance of follow-up practice and repeated exposure to newly learned target vocabulary, absence of which causes the decrease in the delayed post-test scores of the participants. It appears that for long-term retention of new vocabulary items, it is not sufficient to get exposed to the word and its meaning only once (Ko, 2012).

Lastly, according to Watanabe (1997) gloss use itself might hinder retention of vocabulary. Learners who consult glosses might not feel the need to infer meanings from the context, which in turn reduces the amount of processing on the part of the learners. When we consider that " . . . recall is determined by the levels or depth of processing in the encoding process" (Watanabe, 1997, p. 289), it is probable that gloss use somehow hinder vocabulary retention.

In terms of the effectiveness of gloss language on vocabulary retention, there are mixed results in the literature. Laufer and Shmueli (1997) and Arpacı (2016) found that the L1-gloss condition proved more beneficial for retention than the L2-gloss condition. On the other hand, according to the results of Jacobs et al. (1994), Miyasako (2002) Yoshii (2006), and Ko (2012) neither gloss type had a significant influence on vocabulary retention.

Laufer and Shmueli (1997) investigated the relationship between memorization of new words and teaching techniques involving different modes of vocabulary presentation and different languages of glossing. Participants were 128 high school students whose L1 was Hebrew. They were studying in the highest-level English classes. There were five conditions: (i) words presented in isolation, (ii) in one meaningful sentence, (iii) in text-context, (iv) in "elaborated" text context, (v) a list of words (control group) In each condition, except for the control group, 10 words were translated into L1 and the other 10 words were described in L2. The participants in the first two conditions studied the words, whereas the third and fourth groups read a text and took a comprehension test. At the end, participants completed a fill in the blanks type consolidation task. Participants on the control group were asked to check the meanings of the words in their lists and learn them for a quiz. A multiple-choice vocabulary test was administered to all five groups twice,

immediately after the treatment and again five weeks later. Participants other than the control group were not informed about these tests. Results of the study showed that L1-glosses were superior to L2-glosses, for both the immediate and delayed vocabulary tests and irrespective of the method in which the words were learned. Laufer and Shmueli (1997) also stated that the use of a "consolidation task" drew attention to the target words and might therefore contributed to participants' memorization. Thus, Laufer and Shmueli (1997) added, that the kind of learning they had in their study was quasi incidental.

Arpacı (2016) investigated the effectiveness of L1 (Turkish) versus L2 (English) glosses and no gloss condition on incidental vocabulary learning. She administered the delayed post-test 15 days after the immediate test. The test consisted of 16 multiple-choice questions. According to the results of Arpacı (2016), there was a significant decrease in all gloss conditions in their delayed post test results. The mean scores for each gloss group differed significantly from each other: the delayed vocabulary test scores of the participants in the L1-gloss group were the highest, followed by the L2-gloss group, and no gloss group scored the lowest.

Jacobs et al. (1994) administered the delayed post-test four weeks after the immediate test. The test required participants to translate 32 glossed items from Spanish (L2) to English (L1). On the delayed vocabulary test, there were no significant differences between the L1 and L2-gloss groups, or between the gloss and no-gloss groups.

Miyasako (2002) administered the delayed post-test 18 days after the immediate test. The test was in multiple-choice format and consisted of 15 items. The results revealed that there was no significant difference between the delayed

post-test scores of the five gloss conditions, implying that glossing did not affect retention.

Yoshii (2006) administered the delayed post-test two weeks after the immediate test. The test consisted of two sections: a definition-supply and a recognition test. The results revealed that there was no significant difference between glossed conditions. The research design does not include a control group. Therefore, the results of the study did not provide us information as to whether there is a difference between glossed and no gloss conditions.

Ko (2012) administered the delayed post-test four weeks after the immediate test. The test was in multiple-choice format and consisted of 16 items. Although participants in the glossed condition performed significantly better than those in the no gloss condition, no significant difference was observed between L1 and L2-glosses.

Although these results do seem contradictory, we should consider the differences between the procedures followed by each study before deciding whether these findings are comparable or not. For example, the time passed between the immediate and delayed tests was; five weeks for Laufer and Shmueli (1997), two weeks for Yoshii (2006), 15 days for Arpacı (2016), four weeks for Jacobs et al. (1994) and Ko (2012), and 18 days for Miyasako (2002). The test methods used by Yoshii (2006) were a definition-supply test and a recognition test, whereas, Laufer and Shmueli (1997), Arpacı (2016), Miyasako (2002) and Ko (2012) preferred to utilize multiple-choice vocabulary tests. Jacobs et al. (1994), on the other hand, used a translation task.

2.4 The present study

The aim of the present study is to investigate whether access to glosses facilitates incidental vocabulary learning and retention in the L2 (English) and whether there is difference between L1 (Turkish) and L2-glosses in terms of their effects on incidental vocabulary learning, retention and reading comprehension. The participants of the study were randomly assigned to three types of gloss conditions: (i) L1 (Turkish) gloss, (ii) L2 (English) gloss and (iii) no gloss (control group). The research questions of the study are as follows:

- Does access to glosses affect incidental vocabulary learning and reading comprehension in the L2?
- Is there a difference between L1 versus L2-glosses in terms of their effects on incidental vocabulary learning, retention and reading comprehension in the L2?
- Does access to glosses affect the amount of time spent on reading the text?
- Does the type of gloss affect the number of glosses accessed and the amount of time spent on using the glosses?
- Are the frequency of access to glosses and the amount of time spent on glosses related to incidental vocabulary learning and reading comprehension?

CHAPTER 3

METHODOLOGY

3.1 Participants

Participants of the study consist of 74 students (36 female and 38 male) enrolled in an intensive English program at a Turkish state university. According to the results of the placement test that was administered by the university, these students were placed in the upper-intermediate level courses. They take 20 hours of English language courses per week. These courses comprise of: Listening and Speaking (four hours per week), Grammar, Reading and Writing (16 hours per week). At the time of data collection, the students were eight weeks into the spring semester. Along with the 15-week fall semester the students had a total of 460 hours of English language courses. The average age of the participants is 20, and the average age of onset of English language learning is 10.

3.2 Materials and instruments

Materials of the present study include; a reading software, which is used for both reading and testing, a reading text, a comprehension test, three different types of vocabulary tests (form recognition, meaning recognition, and meaning production), a questionnaire and semi-structured interviews with participants who volunteered.

3.2.1 Reading text

A reading text (see Appendix A) was chosen in cooperation with the instructor of the participants. The total number of the words in the text is 493. According to the Flesch-Kincaid readability scale the grade level of the text is measured to be 9 with a

reading ease score of 64. The level of the text was approved by the instructor to be appropriate for the participants' proficiency level. A pilot study was conducted with 20 students who did not take part in the experiment. The students in the pilot study were given the hardcopy version of the text and asked to read and underline the unknown words. As a result, 18 words were underlined as unknown, and all of them were glossed. The word class distribution of these 18 words is as follows: eight nouns, two adjectives, four adverbs, three verbs and one idiom (see Appendix B).

The reading text was digitized, and glosses were incorporated through hypermedia links. The software was developed using Java programming language. The tool is a desktop application, which can be uploaded to a flash disk and later downloaded to any PC or laptop. Each paragraph of the text was presented on a separate screen, with a total of six screens. The glossed words were shown in blue and boldfaced. Upon clicking on a glossed word, the definition of the word either in L1 or in L2 appeared in a box in the middle of the screen (see Figure 1). When the

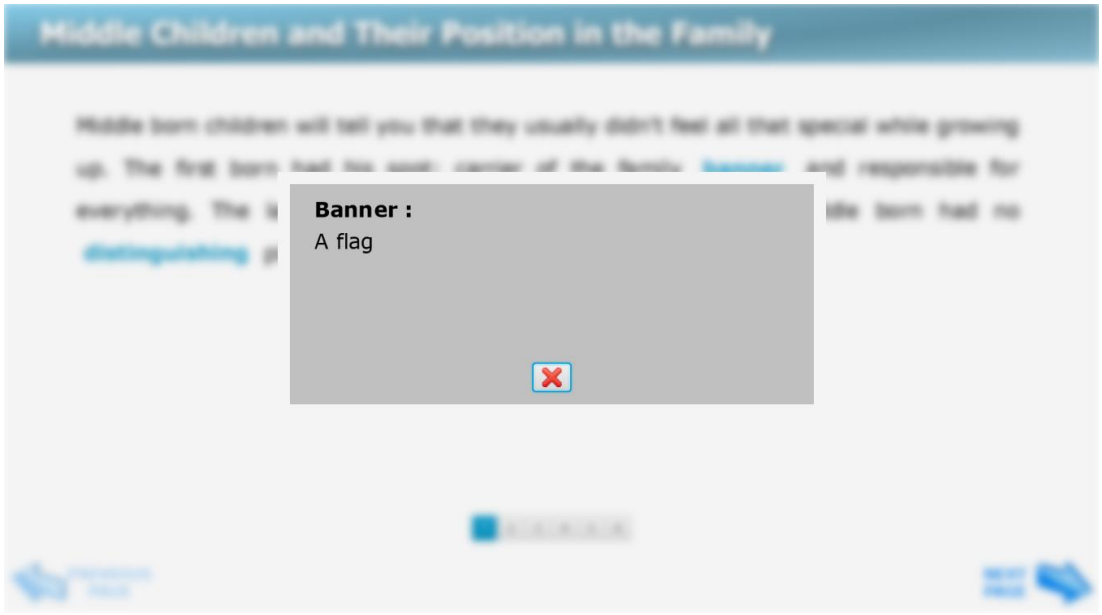


Figure 1. A screenshot showing how glosses are presented in the software

box appeared, the reading text at the background became blurred. This is to make sure to track the exact time spent on glosses by the participants. The participants' interaction with the text was tracked in terms of the frequencies of access to glosses, and the amount of time spent on gloss use and on reading the text.

A navigation map and also two arrow-buttons at the bottom of the screens were provided to navigate back and forth between pages (see Figure 2). The navigation map showed the current page, previous pages and the remaining pages to be read. To navigate across the pages, the participants could either click on the page numbers on the navigation map or on the arrow-buttons to simply go to the next or the previous page (see Figure 2).

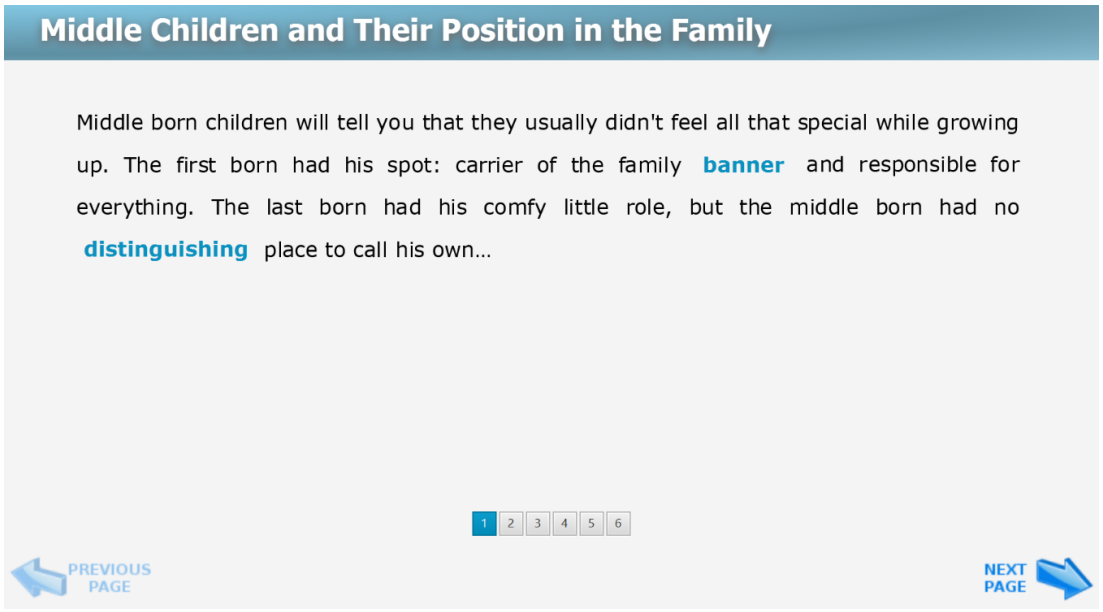


Figure 2. A screenshot showing the screen layout of the reading text screen

The reading text was provided in three different conditions: without glosses, with L1-glosses, and with L2-glosses. Participants were randomly assigned to these three conditions. Thus, the number of participants in the no-gloss condition was 25, in the L1-gloss condition was 25, and in the L2-gloss condition was 24. The software was designed so as to embed the assigned conditions in the participant usernames.

Participants entered their usernames to start the software and having recognized the username the software started with the assigned condition of the relevant participant.

3.2.2 Reading comprehension test

Reading comprehension is not the primary question of the present study. The current study makes use of a comprehension test as a medium to get the participants focus on the meaning of the text; hence induce incidental vocabulary learning. In order to measure incidental vocabulary learning, participants should not be informed that they will be administered a vocabulary test after they read the text. Instead, they were asked to read for meaning, and they were informed that they would be administered a reading comprehension test upon reading the text.

The reading comprehension test consisted of seven multiple choice questions (see Appendix C). All of the questions were literal comprehension questions. Each question had four options. The test was developed by the researcher and expert opinion was taken for each item and also for ensuring the plausibility of the distractors. Each correct answer was given one point and the maximum possible score on the test was seven. The test was administered on-screen after the participants finished reading the test. The text was not accessible while answering the questions. The rationale behind this procedure is based on Pederson's (1986) argument that in order to have a valid measurement of reading comprehension the text should not be available to the students. Before reading the text, the students were informed that they would not have access to the text while answering the comprehension questions. Also, upon finishing reading the text a warning message appeared on the screen to inform the participants that if they continue with the test, they would not be able to access the reading text anymore.

3.2.3 Vocabulary tests

Participants' incidental vocabulary learning was assessed through unannounced vocabulary tests administered after the participants completed the reading comprehension test. There were three vocabulary tests assessing different levels of vocabulary knowledge: form recognition, meaning recognition, and meaning production. All of the tests included 18 target words. Each correct answer was worth one point on all of the tests. The tests were administered on-screen.

Participants received the same meaning production and meaning recognition tests three weeks later, again unannounced, in order to see whether the new vocabulary items were retained or not. The ordering of the questions and the distractors in delayed test were changed in order to control for the possibility of answering correctly due to sight memorization as opposed to answering due to having actually acquired the meaning of the word. This time the tests were administered on paper.

3.2.3.1 Form recognition

The form recognition test included a checklist in which participants were asked to indicate the words they remembered from the text (see Appendix D). There were 26 items consisting of 18 target items and eight fillers. The Cronbach's alpha for the test was .63. The maximum possible score on the test was 18.

3.2.3.2 Meaning production

The test required the participants to write the L1 or L2 equivalents or synonyms of the target words. All of the 18 glossed words were included in the test as target words (see Appendix E). The Cronbach's alpha value for the post-test was .71 and

for the delayed post-test it was .62. The maximum possible score on the test was 18. The tests were scored by two independent raters. The inter-rater reliability was found to be .98 for the post-test, and .99 for the delayed post-test.

3.2.3.3 Meaning recognition

The test involved the matching of the given definitions with the target words (see Appendix F). There were 26 definitions and eight of them were fillers. The maximum possible score on the test was 18. The Cronbach's alpha value for the post-test was .73, and for the delayed post-test it was .66.

3.2.4 Student opinion questionnaire

The questionnaire included questions about students' demographics and their preferences on having or not having glosses in their reading materials. Students are also required to state the reason behind their preferences. Lastly, students were asked to state in which language (L1 or L2) they would prefer the glosses to be (see Appendix G).

3.2.5 Interview

Twenty-four volunteer students who were exposed to glossed texts were given semi-structured interviews after the delayed post-test. Seven interview questions were prepared in advance (see Appendix H). Participants were free to ask questions and make comments. They were asked to talk about their reading experience with glosses. They stated whether they found glosses useful or not, and in what conditions they felt the need to consult the glosses. They shared their preferences of gloss language, and they were also asked to state the reason for these preferences.

Interviews were recorded using a recording device, and later they were transcribed verbatim in preparation for thematic analysis.

3.2.6 Pilot study

The reading text and all of the tests were piloted with 23 upper-intermediate level preparation class students who did not take part in the experiment. It was observed that the average time spent on reading and test taking was 30 minutes in total. In line with the feedback given by the pilot study's participants, definitions of the glossed words and the questions were revised to avoid any confusion.

3.3 Procedures

The data were collected in the computer laboratory of the participants' university in three sessions held in three consecutive days. In each session students were randomly assigned to three conditions. Firstly, the participants were given brief information about the study. The researcher demonstrated how to use the software. Participants were asked to read the text thoroughly to be able to correctly answer the comprehension test that they would have immediately after the reading session. The total time allocated for the procedure was 30 minutes.

Upon reading the text and answering the reading comprehension test, the participants were administered an unannounced (in order to measure incidental vocabulary learning) on-screen vocabulary test. The vocabulary test consisted of three sections: a form recognition test, a meaning production test and a meaning recognition test. After finishing the tests, students were asked to fill out a student opinion questionnaire.

Three weeks later, again unannounced, the same meaning production and meaning recognition tests were administered, in order to see whether the new vocabulary items were retained or not. The ordering of the questions and the distractors were changed in order to control for the possibility of answering correctly due to sight memorization as opposed to answering due to having actually acquired the meaning of the word. After the delayed test, 24 students were interviewed in order to collect data about their opinions on the glossed reading experience that they had. Figure 3 summarizes the data collection procedures of the present study.

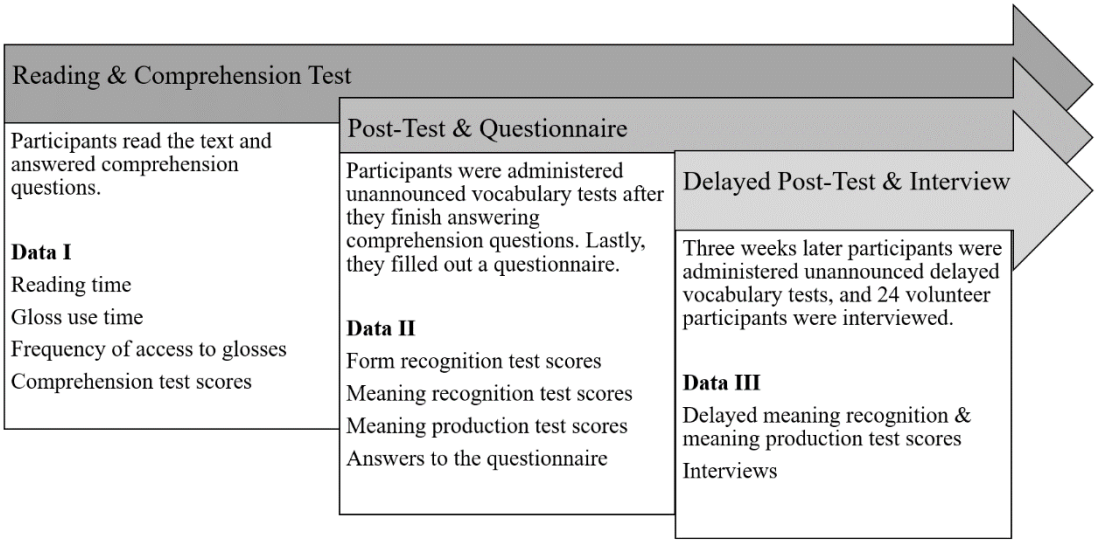


Figure 3. Data collection procedures of the present study

3.4 Data analysis

The reading comprehension test scores were analyzed by conducting a Kruskal-Wallis nonparametric test, because the data were not normally distributed. On the other hand, the scores of the vocabulary tests were analyzed using analyses of variance, because the data were normally distributed. The form recognition test scores were analyzed using one-way ANOVA. For the Meaning Production and the Meaning Recognition test scores a 2 (Type of vocabulary knowledge: meaning

recognition, meaning production) × 2 (Time: Post, Delayed) × 3 (Gloss condition) mixed-design ANOVA was conducted with time and type of vocabulary knowledge as repeated measures factors and condition as between groups factor. An independent samples t-test was conducted in order to see whether total look up frequencies of L1-gloss and L2-gloss conditions differed significantly. A Kruskal-Wallis nonparametric test was conducted to find out whether there is a significant difference between the gloss groups in terms of the total amount of time spent on the reading task. Because total look up frequency data was not normally distributed, Spearman rank order correlation was used in order to analyze the relationship between total look up frequency and reading comprehension test scores as well as vocabulary test scores. The relationship between gloss use time and the test scores were also analyzed using Spearman rank order correlation, because gloss use time data was not normally distributed. In order to analyze the questionnaire data frequency counts and percentages were recorded. Thematic analysis was used to analyze interview data.

CHAPTER 4

RESULTS

4.1 Effects of glosses on reading comprehension

Table 1 provides the descriptive statistics for the reading comprehension test scores of the three groups. There is a small difference between the scores of the participants on three different conditions. The highest mean score belongs to the L1-gloss group followed by the no-gloss group and the L2-gloss group has the lowest mean score.

Table 1. Descriptive Statistics for Reading Comprehension Test Scores

Group	*N*	*M*	*SD*	*Min.*	*Max.*
L1-gloss	25	6.36	0.7	5	7
L2-gloss	24	5.63	1.5	3	7
No gloss	25	6.28	0.74	5	7

Because the data were not normally distributed, a Kruskal-Wallis nonparametric test was conducted on reading comprehension scores in order to see whether the difference among the scores were statistically significant. However, the result was not statistically significant ($H = .286$, 2 *df*, $p > .05$). In other words, the groups did not differ in terms of reading comprehension.

4.2 Incidental vocabulary learning

Of the three tests, the form recognition test was administered only once at post-test. Whereas, meaning recognition and meaning production tests were administered at both post-test and delayed post-test. Therefore, form recognition scores were analyzed separately.

4.2.1 Form recognition

Table 2 provides descriptive statistics for the form recognition test scores of the three groups. The means of the L1-gloss group and the L2-gloss groups are both higher than the No-gloss group on form recognition test.

Table 2. Descriptive Statistics for Form Recognition Test Scores

Group	*N*	*M*	*SD*	*Min.*	*Max.*
L1-gloss	24	15.54	1.67	12	18
L2-gloss	24	14.29	2.27	10	17
No gloss	25	12.76	2.93	8	18

The form recognition test was administered only once at post-test but not at delayed post-test. Therefore, form recognition test scores were analyzed separately. The data were normally distributed. Since the groups were equal in terms of sample size (N = 24), the assumption of equal variances was ignored. A one-way ANOVA on form recognition scores revealed significant differences among the groups ($F_{2,69}$ = 7.536, $p < .01$). Multiple comparisons using the Bonferroni procedure indicated a significant difference between the L1-gloss group and the No-gloss group ($p < .001$). However, the difference between the L2-gloss group and the No-gloss group was not significant ($p > .05$). The difference between the L1-gloss group and the L2-gloss group was not significant ($p > .05$) either.

4.2.2 Meaning production and meaning recognition

Table 3 provides descriptive statistics for the meaning production and meaning recognition test scores of the three groups on post-test and delayed post-test. Although there is some amount of difference in post-test and delayed post-test scores

of all conditions, the decrease in L1-gloss group's performance in the delayed post-test for both meaning production and meaning recognition test is remarkable.

Table 3. Descriptive Statistics of Post-Test and Delayed Post-Test Scores in Meaning Production and Meaning Recognition Tests

	Group	*M*	*SD*
Meaning Production Post-test	L1-gloss	10.24	3.74
	L2-gloss	7.65	1.87
	No gloss	5.26	2.34
Meaning Production Delayed Post-test	L1-gloss	7.6	3.4
	L2-gloss	6.7	2.18
	No gloss	5.87	2.47
Meaning Recognition Post-test	L1-gloss	14.08	2.41
	L2-gloss	10.55	2.48
	No gloss	8.04	2.34
Meaning Recognition Delayed Post-test	L1-gloss	8.08	3.12
	L2-gloss	7.8	2.73
	No gloss	7	2.35

A 2 (Type of vocabulary knowledge: meaning recognition, meaning production) × 2 (Time: Post, Delayed) × 3 (Gloss condition) mixed-design ANOVA was conducted with time and type of vocabulary knowledge as repeated measures factors and condition as between groups factor.[1]

The results revealed a significant main effect of type of vocabulary knowledge [$F(1, 65) = 88.95$, $p < .001$, partial $\eta^2 = .58$] and of time [$F(1, 65) = 88.93$, $p < .001$, partial $\eta^2 = .58$] and of condition [$F(2,65) = 14.44$ $p < .001$ partial $\eta^2 = .31$] . The interaction between vocabulary knowledge and time [$F(1, 65) = 31.21$, $p < .001$, partial $\eta^2 = .32$], as well as the interaction between time and condition [$F(2,$

[1] Normality was checked through histograms, box and whisker plots and normality tests, no severe violations from normality were observed. Box's test of equality of covariance matrices was not significant ($p > .05$) suggesting that the observed covariance matrices of the dependent variable are equal across groups.

65) = 29.85, $p < .001$, partial $\eta^2 = .48$] were also significant. The three way interaction between vocabulary knowledge, time and gloss condition was not significant [$F(2, 65) = 1.92$, $p > .155$]. Nor was the two way interaction between vocabulary knowledge and gloss condition [$F(2, 65) = .09$, $p > .92$].

Figure 4 illustrates the interaction between vocabulary knowledge and time. Follow-up pairwise comparisons through t-tests indicated that, at post-test the participants had significantly higher meaning recognition scores ($M = 10.80$, $SD = 3.46$) than meaning production scores ($M = 7.61$, $SD = 3.46$), $t_{73} = 11.08$, $p < .001$. At delayed test, the scores decreased for both meaning recognition ($M = 7.63$, $SD = 2.76$), $t_{67} = 8.13$, $p < .001$ and meaning production ($M = 6.75$, $SD = 2.83$), $t_{67} = 3.29$, $p < .01$. Although the decrease in meaning recognition was more substantial compared to meaning production, the mean for the former was still significantly higher at delayed test compared to the latter, $t_{67} = 3.18$, $p < .01$.

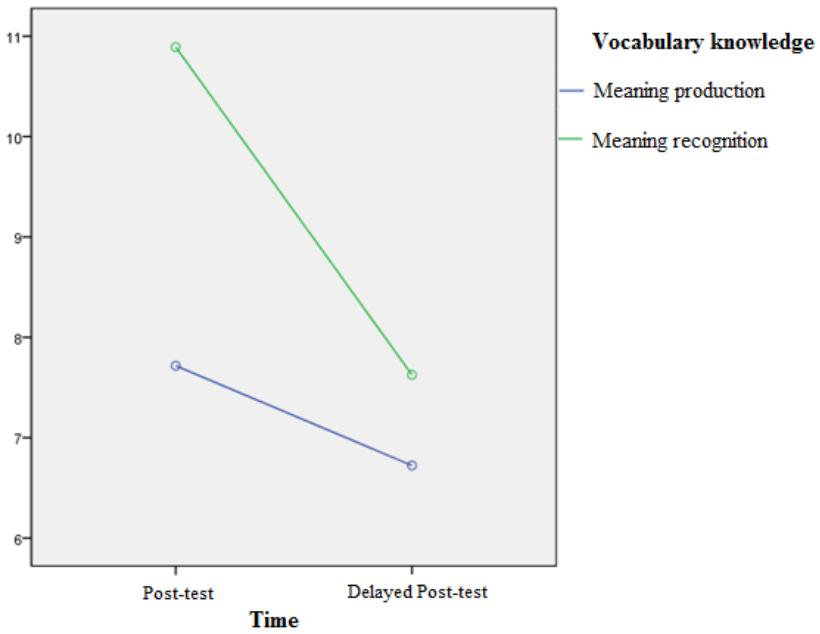


Figure 4. Interaction between vocabulary knowledge and time

Figure 5 illustrates the interaction between gloss condition and time. The result of a univariate analysis of variance on the post-test scores indicated significant differences among the gloss conditions [$F(2, 71) = 34.19$, $p < .001$]. At post-test, the L1-gloss group scored the highest ($M = 12.16$, $SD = .47$) followed by the L2-gloss group ($M = 8.65$, $SD = .48$), and the no gloss group scored the lowest ($M = 6.78$, $SD = .47$).

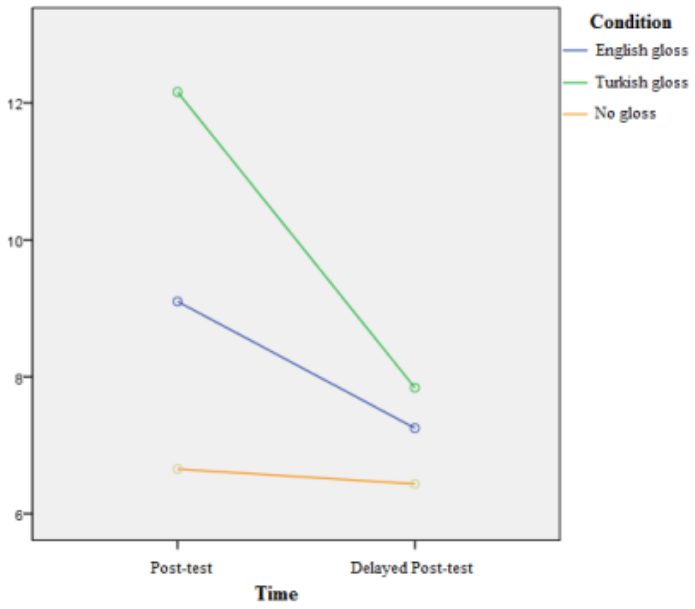


Figure 5 Interaction between gloss condition and time

The result of a univariate analysis of variance on the delayed post-test scores indicated that the gloss conditions did not differ significantly [$F(2, 65) = 1.87$, $p > .16$]. Although the differences between the scores were not significant, the L1-gloss group again scored the highest ($M = 7.84$, $SD = .50$) followed by the L2-gloss group ($M = 7.25$, $SD = .56$) and the no gloss group again scored the lowest ($M = 6.43$, $SD = .53$). The decrease of the L1-gloss group's performance was more substantial compared to the decrease of the L2-gloss group's performance.

4.3 Analyses based on log data

Data regarding the amount of time spent on reading the text and using the glosses as well as the number of times glosses were consulted came from the log files that kept track of the participants' interaction with the text and glosses (see Table 4).

Table 4. Means and Standard Deviations for Text Interaction Measures

	Total Reading Time (min)		Total Look Up Frequency		Gloss Use Time (min)	
Group	*M*	*SD*	*M*	*SD*	*M*	*SD*
English-Gloss	5,92	2,18	14,04	9,02	1,00	0,78
Turkish-Gloss	6,67	2,72	13,80	6,23	0,43	0,20
No Gloss	6,07	1,75	-	-	-	-

An independent samples t-test indicated a nonsignificant difference in the total look up frequencies of L1-gloss condition and L2-gloss condition, $t(47) = .11$, $p > .91$, suggesting that both groups accessed glosses in equal amounts. On the other hand, gloss use time differed significantly with the participants in the L1-gloss condition spending less time using the glosses than those in the L2-gloss condition, $t(47) = 3.48$, $p < .001$. Finally, a Kruskal-Wallis nonparametric test did not point to significant differences across the groups in terms of the total amount of time spent on the reading task ($H = .774$, 2 *df*, $p > .05$). In other words, the groups did not differ in terms of the amount of time they spent on reading the text.

Spearman rank order correlation was used in order to analyze the relationship between total look up frequency and incidental vocabulary learning, because total look up frequency data was not normally distributed. The correlation coefficient resulting from the analysis indicated that there is a moderate positive linear correlation ($r_s = .41$, $p < .01$) between total look up frequency and form recognition test scores. The direction of the relationship is positive, implying that the more

frequently glosses are accessed, the higher scores were gained from the form recognition test. There is also a moderate positive linear correlation ($r_s = .45$, $p < .05$) between total look up frequency and form recognition test scores of the L2-gloss group. Whereas for L1-gloss group there was no significant correlation between total look up frequency and form recognition test scores ($p > .45$). There is a moderate positive linear correlation ($r_s = .50$, $p < .01$) between total look up frequency and meaning production and meaning recognition post-test scores. The direction of the relationship is positive, implying that the more frequently glosses are accessed, the higher scores were gained from the post-test. However, the result of another Spearman rank order correlation revelated that there was no significant correlation between total look up frequency and reading comprehension test scores.

The relationship between gloss use time and test scores were analyzed using Spearman rank order correlation, because the gloss use time data was not normally distributed. There was no significant correlation between gloss use time and reading comprehension test scores. However, a moderate positive linear correlation was observed ($r_s = .32$, $p < .01$) between gloss use time and form recognition test scores. The direction of the relationship is positive, implying that the longer time spent accessing the glosses the higher scores were gained from the form recognition test. Another moderate positive linear correlation ($r_s = .40$, $p < .01$) was observed between gloss use time and meaning production and meaning recognition post-test scores. There was no significant correlation between gloss use time and delayed post-test scores.

4.4 Insights from the questionnaires and interviews

All of the 74 participants were given the questionnaire right after the post-test. 88% of the participants stated that they would prefer to have glosses in their reading materials. Some of the stated reasons for this preference were as follows; (i) glosses are helpful for understanding the text, (ii) they are timesaving, (iii) they are helpful for learning unknown vocabulary, and (iv) glosses are more practical than using dictionaries. The remaining 12% of the participants preferred not to have glosses as part of their reading materials. Some of the stated reasons for this preference were as follows; (i) it is better to figure out the meaning of unknown words by one's own effort, (ii) reading with glosses is confusing, and (iii) dictionaries provide more information than glosses. As for their gloss language preferences, 74% of the participants stated that they would prefer having glosses in L2, whereas 23% of the participants preferred to have glosses in L1, and the remaining 3% of the participants stated that they would prefer having glosses both in L1 and L2.

After the delayed post-test, twenty-four participants volunteered for an interview. They were asked to talk about their reading experience with glosses. They stated whether they found glosses useful or not, and in what conditions they felt the need to consult the glosses. They shared their preferences of gloss language, and they were also asked to state the reason for these preferences.

All of the participants agreed that instant access to the definitions of glossed words facilitated the reading process and accelerated their reading pace. They explained this further by stating that glosses helped them avoid getting interrupted each time they came across an unknown word and try to find its definition. It was also mentioned that without glosses they may have failed to find the meanings of the words that exactly correspond to the given context. On the other hand, two of the

respondents shared their belief that being able to read the text faster than usual caused them to fail in retaining the meanings of the glossed words for future use. This statement is in line with Watanabe's (1997) suggestion that gloss use itself might hinder retention of vocabulary, because learners who consult glosses might not feel the need to infer meanings, which in turn reduces the amount of processing on the part of the learners. When the amount of processing reduces, recall gets negatively affected.

33% of the participants found that having separate screens for each paragraph enabled them to read faster and to comprehend the text better. It was also stated that this computer assisted reading experience was more beneficial in terms of navigating through the pages of the reading text than it is with traditional paper-based reading.

67% of the participants stated that they accessed glosses only if they did not know the meaning of the target word. 33% of the participants stated that they accessed all of the glossed material in order to check the accuracy of their vocabulary knowledge. Aust, Kelley and Roby (1993) also found that when glosses were readily accessible, participants used them excessively to look up even familiar information. Two participants stated that they were driven to click on the glossed words as soon as they notice them. The glossed words were presented in blue and boldfaced, and this was the stated reason for these two participants to feel an urge to access glosses.

Participants mentioned several strategies that they apply in order to figure out the meanings of unknown vocabulary. These strategies include, considering the context, paying attention to word stem or suffixes and prefixes, considering which word class the word belongs to. Participants stated that they also paid attention to neighboring words and in a wider scope they may consider looking for clues from

previous or following sentences. They stated that if these strategies were helpful in sorting out the meaning of a glossed word, they did not access glosses.

40% of the participants stated that having accessed the glosses helped them in the delayed post-test. Some of these participants added that they have photographic memory, some others mentioned being a visual learner, both of which enabled them in the delayed post-test to visualize on their minds; the glosses, their position on the computer screen and even their colors. Another 40% of the participants stated that having accessed the glosses did not help them in the delayed post-test. They stated that although they find it useful to access the meaning of unknown words immediately and without interruption, it did not help them to retain the meanings. They also added that they would have retained the meanings if they had searched for them and looked them up from the dictionary. The remaining 20% of the participants stated that having accessed the glosses during reading, partially helped them in the delayed post-test.

Participants were asked whether they would like to have any additional information as gloss material. 50% of the participants stated that they would like also to see an example sentence where the target word is used. They believed that this would have helped them comprehend and retain the meaning of the target word much better. They stated that if they cannot understand the given description of the glossed word, an example sentence may help them to work out the meaning of the unknown word. They stated that making a mental effort to figure out the meaning of the target word helps them retain the meaning of the word. 17% of the participants stated that they do not need any additional information than the meaning of the word. They believed that any additional information would distract them and cause them to get disconnected from the text they read. Another 17% of the participants stated that they

would like to have visual material in addition to the meaning of the word. They believed that visual aids are important for them to retain the meaning of the words. The remaining 16% of the participants stated that they would like to have additional information on the pronunciation of the target word.

Participants were asked about whether they would like to read an on-screen text with glosses or a text on paper. 50% of the participants stated that they prefer reading from hardcopy reading material. They stated that they cannot concentrate well when they read on-screen texts. They added that they like to physically touch the material, underline parts of the text that they find important and take notes on the margins of the paper. In addition, they stated that on-screen texts with glosses are very useful and timesaving. They stated that these advantages of on-screen texts may help them get used to it in time. Other 50% of the participants stated that they prefer reading on-screen texts with glosses. They think that it provides a faster, more useful, more comfortable and easier reading experience than reading hardcopy texts. They also stated that glosses make it easier to access the meaning of the unknown vocabulary. However, they added that this ease of access raises a concern that having spent close to no effort for figuring out the meaning of a target word may cause failure in retention of the vocabulary in the long run.

Participants were also asked about their preferred gloss language. 41% of the participants stated that they prefer glosses to be in English. They stated that sometimes L1 translations do not fully correspond to what is meant by the English word. The participants stated also that with L2-glosses, it is also possible to see and learn about the synonyms of the target word. They mentioned that when they read a text in English, they somehow switch to English language mode (that they start to think in English) and that they do not want to have any connection with Turkish

language. They stated that coming across a gloss definition in Turkish while reading a text in English confused them, and that they found it hard to fit the meaning in place. 23% of the participants stated that they prefer glosses to be in Turkish. They stated that they are more comfortable with Turkish, and that it helps them fully comprehend the meaning. The remaining 36% of the participants stated that they would like to have glosses in both English and Turkish. They stated that they think L2- glosses are helpful to acquire the meaning of the target word. Whereas, they believed that L1-glosses are helpful for fully comprehending a text. They also shared their concern that if glosses are in English, there may be unknown words in the definition, too, making gloss use pointless. In some situations, L1 translations may not fully correspond to what is meant by the English word. In order to avoid such situations, the participants proposed, it is better to have glosses in both English and Turkish.

CHAPTER 5

DISCUSSION AND CONCLUSION

Findings of this study suggest that gloss use did not facilitate reading comprehension irrespective of the gloss condition, which confirms the findings of previous studies (Johnson, 1982; Pak, 1986; Jacobs, Dufon, & Hong, 1994; Davis & Lyman-Hager, 1997; Miyasako, 2002; Ariew & Ercetin, 2004; Akbulut, 2007; Al-Jabri, 2009; Cheng & Good, 2009), pointing to lack of a significant effect of text based glosses on reading comprehension. On the other hand, findings of Jacobs (1994) and Arpacı (2016) suggested that gloss use facilitated reading comprehension. As Jacobs (1994) argues, effectiveness of glosses on reading comprehension might be influenced by a number of different factors such as: the difficulty level of the text, the proficiency level of the readers, the items that are chosen to be glossed, and the percentage of items that are glossed in the text. In addition, Ko (2005) suggests that while deciding on which words to gloss, the words' relative importance in understanding the text should be considered. Another suggestion made by Ko (2005) is that in order to get more reliable results there should be consistency in terms of how reading comprehension gets assessed.

Findings of the current study revealed that use of L1-glosses was more effective than no gloss condition in incidental vocabulary learning as measured by the form recognition test. Another finding is that both L1 and L2-glosses were effective in incidental vocabulary learning as measured by the meaning production and meaning recognition tests, with L1-glosses yielding better test scores than L2-glosses did.

At post-test the participants had significantly higher meaning recognition scores than meaning production scores. This result is consistent with the finding of Abraham's (2008) meta-analysis suggesting that there is a comparatively larger effect for receptive tests than for productive tests. In addition, the L1-gloss group scored the highest followed by the L2-gloss group, and the no gloss group scored the lowest, which confirms Arpacı's (2016) findings, showing significant differences between glossed groups and no gloss group, and between the vocabulary test scores of L1 and L2-gloss groups, favoring L1-gloss condition. The poor performance of participants exposed to the no gloss condition may be explained by ineffective guessing strategies for figuring out the meanings of unfamiliar words from context. As mentioned earlier, in order to be able to use guessing strategies effectively, learners need to know the meaning of 95–98% of the words in a text (Nation & Meara, 2010). Due to their proficiency level (upper-intermediate for the current study and lower-intermediate for Arpacı (2016)) the participants had a rather limited overall L2 knowledge in terms both of vocabulary and guessing strategies. Accordingly, as was also shown by Ko (2012), L1-glosses were more effective for learners with lower-proficiency levels. Schmitt (2008) suggested that as long as the learner can understand the meaning, the language of the glosses does not matter, suggesting that L1-glosses can be used with less-proficient learners.

Another factor to explain why glossed conditions were better at incidental vocabulary learning might be the simple fact that glosses were boldfaced. This alone directs participants' (the ones in gloss conditions) attention to the glossed vocabulary and when attention is paid noticing takes place. Schmidt (1990) suggested that in order for input to turn into intake noticing is necessary and sufficient. If glossed words were not boldfaced, participants may have ignored them as part of an

avoidance strategy that is used when a word is not deemed crucial for general comprehension (Schmitt, 1997).

Findings of the study suggest that gloss condition does not affect vocabulary retention. These findings are in line with the results of previous studies (Jacobs et al., 1994; Miyasako, 2002; Yoshii, 2006; Ko, 2012) which also suggested that neither type of gloss had a significant influence on vocabulary retention. For the current study, although the differences between the scores were not significant, the L1-gloss group scored the highest followed by the L2-gloss group and the no gloss group again scored the lowest on the delayed post-test. Research results of Laufer and Shmueli (1997) and Arpacı (2016) suggested that the L1-gloss condition proved more beneficial for retention than the L2-gloss condition. However, in their study Laufer and Shmueli (1997) used an additional "consolidation task" right after the reading phase. This task drew participants' attention to the target words and also gave them an additional chance to practice the target vocabulary. Therefore, "consolidation task" might have contributed to participants' memorization. Findings of Arpacı (2016) also suggested that L1-gloss condition was superior to L2-gloss condition in terms of vocabulary retention. The time passed between the immediate and delayed tests was 15 days for Arpacı (2016), whereas for the current study it was three weeks. The passage of more time might have resulted in failure of recall on the part of the current study's participants.

At delayed post-test, the scores decreased for both meaning recognition and meaning production tests. Although the decrease in meaning recognition was more substantial compared to meaning production, the mean for the former was still significantly higher at delayed test compared to the latter. This can be explained by Laufer and Paribakht's (1998) research results suggesting that passive (receptive)

vocabulary of L2 learners is larger than their active (productive) vocabulary. The finding that the decrease in meaning recognition was more substantial compared to meaning production can be interpreted as passive (receptive) vocabulary's vulnerability to passage of time in terms of memory performance. Active (productive) vocabulary, on the other hand, appears to have a more persistent trace and hence is less vulnerable to passage of time.

The decrease in L1-gloss group's performance in the delayed post-test for both meaning production and meaning recognition test is remarkable. This finding may suggest that although L1-glosses have a potentially positive effect on incidental vocabulary learning, the vocabulary gains through L1-glosses need to be reinforced in order to be retained (Jacobs et al.,1994). Schmitt (2008) also stated that in order to improve incidental vocabulary learning it should be reinforced afterwards with intentional learning tasks.

Findings of this study suggest that L1 and L2-gloss condition groups accessed glosses in equal amounts. In other words, type of gloss did not affect the number of glosses accessed. However, gloss use time differed significantly with the participants in the L1-gloss condition spending less time using the glosses than those in the L2-gloss condition. This is probably due to participants' processing efficiency in the L1. As suggested by the revised hierarchical model (Kroll & Stewart, 1994), in earlier stages of L2 acquisition, learners associate L2 words with their L1 translations in order to access the meaning that is already represented for those concepts. This might explain the reason why the participants in the L1-gloss condition spent less time accessing the glosses. Finally, the findings suggested that access to glosses does not affect the amount of time spent on reading the text. Türk and Erçetin (2014) also reported a similar finding. The results of their study suggested that there was not a

significant difference between the two gloss conditions (interactive versus simultaneous display of multimedia glosses) in terms of time spent on reading the text.

Within the L2-gloss group the more frequently glosses were accessed, the higher scores were gained from the form recognition test. However, findings of the study also suggested that, form recognition test scores of the L1-gloss group were higher than that of the L2-gloss group.

The more frequently glosses were accessed, the higher scores were gained from all post-tests (form recognition, meaning production, meaning recognition) irrespective of gloss type. This finding suggests that learners should be encouraged to access glosses for better vocabulary gains. Learners might not access glosses thinking that they already know the meaning of the target word, or that they can correctly infer its meaning depending on the context, however it is possible that they might be mistaken. Incidental vocabulary learning was also found to be positively related to the time spent accessing the glosses. However, gloss use time and delayed post-test scores as well as reading comprehension test scores were not found to be related.

A positive relationship was found between L2-gloss group's frequency of access to glosses and their form recognition test scores. Total look up frequency and reading comprehension test scores as well as delayed post-test scores were not found to be related.

According to the qualitative data collected by means of a questionnaire, 74% of the participants of the current study stated that they would prefer having glosses in L2, whereas 23% of the participants preferred to have glosses in L1, and the remaining 3% of the participants stated that they would prefer having glosses both in

L1 and L2. Literature has different findings in terms of preferred gloss languages of the participants of different studies. There are some studies in which the majority of the participants' preferred gloss language was observed to be their L1 (Laufer & Hill, 2000; Bell and LeBlanc, 2000; Arpacı, 2016), whereas in some others it is their L2 (Ko, 2012 ; Jacobs et al.,1994). In Azari (2012), 55.26% of the participants preferred to have both L1 and L2 (Persian and English) glosses, 25.1% preferred to have L1 (Persian) glosses, and 14.47% opted for having L2-glosses. Taylor (2006) suggests that learners whose proficiency level is intermediate and above feel comfortable consulting L2-glosses. The present study reveals a striking contrast: although L1-glosses were found to be more effective than L2-glosses in terms of incidental vocabulary learning, majority of the participants stated that they would prefer having glosses in L2. This is to say, participants have misleading perceptions about the use of L1 in L2 learning. The results of the study may help learners and teachers alike overcome such bias and benefit from the effectiveness of L1-glosses.

Participants who were interviewed stated that glosses assisted their comprehension: a finding that is also reported by many other studies (Arpacı, 2016; Jacobs et al., 1994; Ercetin, 2003; Ko, 2005; Akbulut, 2007; Cheng and Good, 2009; Ko, 2012). Participants of the present study also perceived the computerized glosses to be useful and timesaving, some other studies also reported findings in line with these perceptions (Davis & Lyman-Hager, 1997; Ariew & Erçetin, 2004).

5.1 Pedagogical implications

Findings of the present study may help instructors and reading software developers to focus closely on the choices and needs of language learners, hence create more individualized reading software. According to the results of the current study, gloss

use found to be effective in incidental vocabulary learning. Material designers for CALL environments may consider using glossing wherever possible. L1-glosses may especially be preferred for learners whose proficiency level is upper-intermediate and below. The findings indicating that incidental vocabulary gains diminish in time implies that incidental vocabulary learning should be reinforced afterwards with intentional tasks that provide learners opportunities for deeper and more elaborate processing of target words for better retention. The present study also revealed that the participants were biased against the use of L1-glosses. The results of the study may help learners and teachers alike overcome such misleading perceptions about the use of L1-glosses and benefit from the effectiveness of L1-glosses on incidental vocabulary learning.

5.2 Limitations and suggestions for further research

The participants of the present study consisted of upper-intermediate level English language learners. Further research concerning participants with different proficiency levels is needed to better understand the role of language proficiency on the effectiveness of gloss language on incidental vocabulary learning, retention, and reading comprehension.

Another possible limitation is the sample size and the sampling method. Total number of participants is 74 and the sampling method is convenience sampling. The proficiency level of the participants in the current study was not determined by using standardized testing instruments. It was determined by the placement test that was administered by the university. Hence, the proficiency level is arbitrary. Therefore, the results cannot be generalized to all upper-intermediate level English language learners.

APPENDIX A

READING TEXT

MIDDLE CHILDREN AND THEIR POSITION IN THE FAMILY

Middle born children will tell you that they usually didn't feel all that special while growing up. The first born had his spot- carrier of the family **banner** and responsible for everything. The last born had his comfy little role, but the middle born had no **distinctive** place to call his own…

Middle-borns just seem to be easily **overlooked**, and maybe that's why there are so few pictures of them in the family photo album. There may be hundreds, **seemingly** thousands, of pictures of the firstborn. For some strange reason, however, which I have confirmed by **polling** middle-born children around the world, there are **seldom** many pictures of the middle child, and what photos there are have him included with the others—squeezed again between the older sibling and the younger sibling.

Another thing that can be said of many middle-born children is that they typically place great importance on their peer group. The middle child is well known for going outside the home to make friends faster than anybody else in the family. When a child feels like a **fifth wheel** at home, friends become very important; as a result, many middle children (but not all, of course) tend to be the social lions of the family. While firstborns, typically, have fewer friends, middle children often have many.

Middle children have a **propensity** to leave home first and live **farther** away from the family than anyone else. I observed a dramatic **illustration** of this tendency

while I was a guest on Oprah Winfrey's show. The subject that day was **sibling rivalry**. Three charming young women, all sisters, were among the guests, and we quickly learned that the firstborn and the last born were residents of the Eastern state where they had grown up. They had settled down near their parents and other family members. But the middle child had moved to the West Coast.

I suppose she could have gotten another two thousand miles farther away by moving to Hawaii, but her point was still well made. Middle children are the ones who will most often physically distance themselves from the rest of the family. It's not necessarily because they're **on the outs with** everyone else. They simply like to do their own thing, make their own friends, and live their own lives.

All of this is not to say that middle children totally **ignore** their siblings or the rest of the family. One common characteristic of the middle child is that she is a good mediator or **negotiator**. She comes naturally into this role because she's often right in the middle, between big brother and little sister, whatever the case may be. And because she can't have Mom or Dad all to herself, she learns the fine art of **compromise**. **Obviously**, these skills are **assets** in adult life, and middle children often become the **best adjusted** adults in the family.

APPENDIX B

GLOSSED WORDS

banner (n)

distinctive (adj)

seemingly (adv)

poll (v)

seldom(adv)

fifth wheel (n)

propensity (n)

farther (adv)

illustration (n)

sibling rivalry (n)

on the outs (with somebody)

ignore (v)

negotiator (n)

compromise (n)

obviously (adv)

asset (n)

best adjusted (adj)

overlook (v)

APPENDIX C

COMPREHENSION TEST

Section 1. Please choose the correct answer.

1. Which one of the following is NOT a characteristic of middle children?

A) Outgoing

B) Negotiator

C) Makes compromises

D) Has fewer friends

2. Which one of the following is NOT a disadvantage of being a middle child?

A) being easily overlooked

B) having no distinctive place in the family

C) feeling like a fifth wheel at home

D) learning the fine art of compromise

3. According to the passage middle-born children _____.

A) are responsible for everything

B) are the carrier of the family banner

C) leave home first

D) have many pictures in the family photo album

4. Middle-born children physically distance themselves from the rest of the family because _____.

A) they like to do their own thing and live their own lives.

B) they are on the outs with everyone else in the family.

C) they totally ignore their siblings and the rest of the family.

D) they are the social lions of the family.

5. According to the passage middle-born children place great importance on their _____.

A) parents

B) siblings

C) peer group

D) business partners

6. According to the passage middle children often become the best adjusted adults in the family because _____.

A) they are squeezed between the older sibling and the younger sibling.

B) they live farther away from the family than anyone else.

C) they attract their parents' attention quite easily.

D) they learn how to make compromises and they become good negotiators.

7. According to the passage, if you talk to middle born children, they will tell you that _____.

A) they had thousands of pictures of them in the family photo album.

B) they usually didn't feel all that special while growing up.

C) they didn't feel like a fifth wheel at home.

D) they had settled down near their parents and other family members.

APPENDIX D

FORM RECOGNITION TEST

Section 2. Please put a check mark ✓ on the box next to each word/ phrase that you saw in the text.

banner	□	on the outs with	□
distinctive	□	obscure	□
drought	□	ignore	□
promise	□	negotiator	□
overlook	□	scribble	□
seemingly	□	compromise	□
poll	□	outlook	□
sob	□	obviously	□
seldom	□	asset	□
fifth wheel	□	best adjusted	□
wisp	□	sibling rivalry	□
propensity	□	illustration	□
farther	□	pool	□

APPENDIX E

MEANING PRODUCTION TEST

Section 3. Please write down the meanings of the following words either in English or in Turkish to the space provided.

banner	
distinctive	
overlook	
seemingly	
poll	
seldom	
fifth wheel	
propensity	
farther	
illustration	
sibling rivalry	
on the outs (with somebody)	
ignore	
negotiator	
compromise	
obviously	
asset	
best adjusted	

APPENDIX F

MEANING RECOGNITION TEST

Section 4. Match the following words from column A with their definitions in column B.

Column A	Column B
1 _ _ banner (n)	a. very rarely or almost never
2 ___ distinctive (adj)	b. appearing to have a particular quality, when this may or may not be true
3 ___ seemingly (adv)	c. to deliberately pay no attention to something that you have been told or that you know about
4 ___ poll (v)	d. a flag
5 ___ seldom (adv)	e. very often or many times
6 ___ fifth wheel (n)	f. emotionally healthy and able to deal well with the problems of life, superlative form of "well adjusted"
7 ___ propensity (n)	g. having a special quality, character, or appearance that is different and easy to recognize.
8 ___ farther (adv)	h. arguing or not agreeing with someone
9 ___ illustration (n)	i. a shorter distance
10___ sibling rivalry (n)	j. to ask a lot of people the same questions in order to find out what they think about a subject
11 ___ on the outs (with somebody)	k. a greater distance than before or than something else; a comparative form of 'far'
12 ___ ignore (v)	l. to not notice something, or not see how important it is
13 ___ negotiator (n)	m. a natural tendency to behave in a particular way
14 ___ compromise (n)	n. someone who is not wanted in a particular group of people

5___ obviously (adv)	o. something or someone that is useful because they help you succeed or deal with problems
16 ___ asset (n)	p. a statement that you will definitely do or provide something or that something will definitely happen
17 ___ best adjusted (adj)	q. the things that a company owns, that can be sold to pay debts
18 ___ overlook (v)	r. to combine your money, ideas, skills etc. with those of other people so that you can all use them
	s. a story, event, action etc. that shows the truth or existence of something very clearly
	t. how likely something is, sometimes calculated in a mathematical way
	u. competition between brothers and sisters for their parents' attention or love
	v. not well known and usually not very important
	w. an agreement that is achieved after everyone involved accepts less than what they wanted at first, or the act of making this agreement
	x. not officially allowed to meet, exist, or be used
	y. someone who takes part in official discussions, especially in business or politics, in order to try and reach an agreement
	z. clearly

APPENDIX G

STUDENT OPINION QUESTIONNAIRE

Student Number	
Age	
Gender	☐ Female ☐ Male
When did you start to learn English?	

Gloss: The definitions or translations of unknown words presented at the end of a text or near the unknown words. Computerized glosses allow us to reach the definitions by clicking on the unknown words.

1. Would you prefer to read a text with or without gloss? Why?

..... because .....

.....

.....

2. In which language would you prefer the glosses to be?

☐ Turkish ☐ English

APPENDIX H

INTERVIEW QUESTIONS

1. How did you find this computer mediated reading experience? Was it any different from the way you used to do your readings?

2. What do you think about the difficulty level of the text? Why?

3. Did glosses facilitate the reading process? How?

4. When did you consult glosses?

5. Other than glosses did you use any strategies to figure out the meaning of unfamiliar vocabulary?

6. Do you think having accessed the glosses helped you in the delayed post-test?

7. Would you like to have any additional information as gloss content?

REFERENCES

Abraham, L. B. (2008). Computer-mediated glosses in second language reading comprehension and vocabulary learning: A meta-analysis. *Computer Assisted Language Learning, 21*(3), 199-226.

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.

Al-Jabri, S. (2009). The effects of L1 and L2 glosses on reading comprehension and recalling ideas by Saudi students. *Umm Al-Qura University Journal of Social Sciences, 1*, 11-27.

Ariew, R., & Ercetin, G. (2004). Exploring the potential of hypermedia annotations for second language reading. *Computer Assisted Language Learning, 17*(2), 237-259.

Arpacı, D. (2016). The effects of accessing L1 versus L2 definitional glosses on L2 learners' reading comprehension and vocabulary learning. *Eurasian Journal of Applied Linguistics, 2*(1), 15-29.

Aust, R., Kelley, M. J., & Roby, W. (1993). The use of hyper-reference and conventional dictionaries. *Educational Technology Research and Development, 41*(4), 63-74.

Azari, F. (2012). Gloss or no gloss? EFL learners' preference. *American International Journal of Social Science, 1*(2), 75-85.

Bell, F. L., & LeBlanc, L. B. (2000). The language of glosses in L2 reading on computer: Learners' preferences. *Hispania, 83*(2), 274-285.

Chen, H. (2002). Investigating the effects of L1 and L2 glosses on foreign language reading comprehension and vocabulary retention. Paper presented at the annual meeting of the Computer-Assisted Language Instruction Consortium, Davis, CA.

Cheng, Y. H., & Good, R. L. (2009). L1 glosses: Effects on EFL learners' reading comprehension and vocabulary retention. *Reading in a Foreign Language*, *21*(2), 119-142.

Collentine, J. (2000). Insights into the construction of grammatical knowledge provided by user-behavior tracking technologies. *Language Learning & Technology*, *3*(2), 44-57.

Chun, D. M. (2006). CALL technologies for L2 reading. In L. Ducate & N. Arnold (Eds.), *Calling on CALL: From theory and research to new directions in foreign language teaching,* (pp. 69–98). CALICO Monograph Series Volume 5. Texas State University. San Marcos, TX: CALICO Publications.

Davis, J. N. (1989). Facilitating effects of marginal glosses on foreign language reading. *The Modern Language Journal*, *73*(1), 41-48.

Davis, K. A. (1995). Qualitative theory and methods in applied linguistics research. *TESOL Quarterly*, *29*(3), 427-453.

Davis, J. N., & Lyman-Hager, M. A. (1997). Computers and L2 reading: Student performance, student attitudes. *Foreign Language Annals*, *30*(1), 58-72.

Ercetin, G. (2003). Exploring ESL learners' use of hypermedia reading glosses. *CALICO Journal*, *20*(2), 261-283.

Hashemian, M., & Fadaei, B. (2013). The effect of lexical glossing types on Persian L2 learners' reading and listening. *Procedia-Social and Behavioral Sciences*, *70*, 490-500.

Herman, P. A., Anderson, R. C., Pearson, P. D., & Nagy, W. E. (1987). Incidental acquisition of word meaning from expositions with varied text features. *Reading Research Quarterly*, *22*(3), 263-284.

Hulstijn, J. H. (2001). Intentional and incidental second-language vocabulary learning: A reappraisal of elaboration, rehearsal and automaticity. In P. Robinson (Ed.), *Cognition and second language instruction*, (pp. 258-286). Cambridge: Cambridge University Press.

Jacobs, G. M. (1994). What lurks in the margin: Use of vocabulary glosses as a strategy in second language reading. *Issues in Applied Linguistics*, *5*(1), 115-137.

Jacobs, G. M., Dufon, P., & Hong, F. C. (1994). L1 and L2 vocabulary glosses in L2 reading passages: Their effectiveness for increasing comprehension and vocabulary knowledge. *Journal of Research in Reading*, *17*(1), 19-28.

Johnson, P. (1982). Effects on reading comprehension of building background knowledge. *TESOL Quarterly*, *16*(4), 503-516.

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.

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), 149-174.

Kroll, J. F., & Sunderman, G. (2003). Cognitive processes in second language learners and bilinguals: The development of lexical and conceptual representations. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition*, (pp. 104-129). Malden, MA: Blackwell.

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., & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics*, *22*(1), 1-26.

Laufer, B., & Paribakht, T. S. (1998). The relationship between passive and active vocabularies: Effects of language learning context. *Language Learning, 48*(3), 365-391.

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

Lomicka, L. (1998). "To gloss or not to gloss": An investigation of reading comprehension online. *Language Learning & Technology*, *1*(2), 41-50.

Lyman-Hager, M. A., & Davis, J. F. (1996). The case for computer-mediated reading: Une Vie de Boy. *French Review*, *69*(5), 775-792.

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*(4), 469-479.

Nation, P. (2003). The role of the first language in foreign language learning. *Asian EFL Journal*, *5*(2), 1-8.

Nation, P., & Meara, P. (2010). Vocabulary. In N. Schmitt (Ed.), *An introduction to applied linguistics*, (pp. 34-52). London: Hodder Education.

Pak, J. (1986). *The effect of vocabulary glossing on ESL reading comprehension.* Unpublished manuscript, University of Hawaii, Manoa.

Pederson, K. M. (1986). An experiment in computer-assisted second-language reading. *The Modern Language Journal*, *70*(1), 36-41.

Pitts, M., White, H., & Krashen, S. (1989). Acquiring second language vocabulary through reading: A replication of the Clockwork Orange study using second language acquirers. *Reading in a Foreign Language*, *5*(2), 271-275

Rott, S., & William, J. (2003). Making form-meaning connection while reading: A qualitative analysis of word processing. *Reading in a Foreign Language, 15*, 45-75.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, *11*(2), 129-158.

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

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

Schmitt, N. (2008). Instructed second language vocabulary learning. *Language Teaching Research*, *12*(3), 329-363.

Schmitt, N. (1997). Vocabulary learning strategies. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy*, (pp. 198–227). Cambridge: Cambridge University Press.

Taylor, A. (2006). Factors associated with glossing: Comments on Ko (2005). *Reading in a Foreign Language*, *18*(1), 72-73.

Türk, E., & Erçetin, G. (2014). Effects of interactive versus simultaneous display of multimedia glosses on L2 reading comprehension and incidental vocabulary learning. *Computer Assisted Language Learning*, *27*(1), 1-25.

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.

Wodinsky, M., & Nation, P. (1988). Learning from graded readers. *Reading in a Foreign Language*, *5*(1), 155-161.

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