6-1-2013

The effects of marginal glosses and online dictionary use on incidental receptive and productive vocabulary acquisition through reading

Saliha Arseven

Follow this and additional works at: https://fount.aucegypt.edu/etds

The American University in Cairo

School of Humanities and Social Sciences

The Effects of Marginal Glosses and Online Dictionary Use on Incidental Receptive and Productive Vocabulary Acquisition through Reading

A Thesis Submitted to

the Department of Teaching English to Speakers of Other Languages

in Partial Fulfillment of the Requirements

For the degree of Master of Arts

By

Saliha Arseven

Under the supervision of **Dr. Lori Fredricks**

May/2013

The American University in Cairo

The Effects of Marginal Glosses and Online Dictionary Use on Incidental Receptive and Productive Vocabulary Acquisition through Reading

A Thesis Submitted by

Saliha Arseven

To the Department of Teaching English to Speakers of Other Languages

May/2013

In partial fulfillment of the requirements
For the degree of Master of Arts

Has been approved by

Dr. Lori Fredricks
Thesis Committee Advisor_____
Affiliation_____

Dr. Amira Agameya
Thesis Committee Reader_____
Affiliation_____

Dr. Robert Williams
Thesis Committee Reader_____
Affiliation_____

_____	_____	_____	_____
Dept. Chair	Date	Dean of HUSS	Date

Acknowledgements

I would like to extend my sincerest thanks and gratitude to

My adviser, Dr. Lori Fredricks,

for her invaluable input and guidance

My second reader, Dr. Amira Agameya & my third reader, Dr. Robert Williams

for their extensive feedback and support

My family,

Ayse, Sinan, Elif & Esra,

for always standing by me in all my endeavors

Gulsevin,

for her great support

Without their help and encouragement, this thesis would never have been accomplished.

Thank you!

Abstract

This study tests the Involvement Load Hypothesis (Hulstijn and Laufer, 2001), which suggests that incidental vocabulary acquisition depends on the involvement load (i.e. the amount of mental effort needed to complete a task) required by a task. The premise of this theory is that the use of dictionaries during reading tasks results in better vocabulary acquisition and retention than marginal glosses as the former requires more effort on the part of learners. Fifty-seven Egyptian upper intermediate ESL learners participated in this study. Half of the students were given access to an online monolingual dictionary while completing a reading comprehension task whereas the other half were provided with marginal L2 glosses. Both groups were given four uninformed immediate and four delayed vocabulary tests that measured (1) receptive knowledge of meaning, (2) productive knowledge of meaning, (3) receptive knowledge of word class, and (4) productive knowledge of word class. The results showed a superiority of the online dictionary group in productive tests while no statistically significant difference was seen between the performances of the two groups in receptive tests.

Table of Contents

List of Tables and Figures

Chapter I

1. Introduction

L2 vocabulary acquisition is frequently referred to as a *sine qua non* of improving other L2 skills (Folse, 2006; Hulstijn & Luafer, 2001a; Webb, 2005; Xu, 2010a). Given this fundamental role that vocabulary learning plays in L2 proficiency, it is not surprising that more effective ways of vocabulary acquisition have been investigated by researchers. While the traditional way of vocabulary learning through the memorization of long word lists was practiced for many years, vocabulary acquisition is now widely considered as an ongoing and integrated process within the framework of language learning rather than a self-contained skill that should be taught independently and, oftentimes, explicitly. This idea of vocabulary learning through an integration of skills has been a focus in the literature for the last 30 years and has been shown to be an effective practice (Ahmed, 2011; Bruton, Lopez, & Mesa, 2011; Hulstijn & Laufer, 2001a). Research has shown that much vocabulary is acquired by learners during other L2 tasks even when they have no intention to learn new words, which is called incidental vocabulary acquisition (Ellis, 1994; Gass, 1999; Laufer & Hill, 2003; Schmidt, 1993).

Even though the term "incidental" suggests that there is a "lack of intention" (Hulstijn, Hollander, & Greidanus, 1996), this only refers to the lack of intention for learning new words on the part of learners. For teachers and researchers, there is nearly always a quest as to how an ideal environment for incidental vocabulary acquisition (IVA, henceforth) can be provided to learners. Reading tasks, which are usually considered as the most fruitful environment for IVA (Krashen, 1989; Nation, 2001; Webb, 2005), are most of the time manipulated or modified in a way that promotes IVA.

Glossing, which is a very common way of manipulating reading passages (Ko, 2012), has been shown to be an effective tool in facilitating IVA (Cheng, 2009; Jacobs, Dufon & Hong, 1994; Farvardin & Biria, 2012; Ko, 2012; Xu, 2010a). In general terms, glossing can be defined as providing short descriptions or synonyms – L1 or L2 – of important words in a reading passage. Nowadays, various forms of glossing such as L1-L2 glossing, picture glossing, MCQ glossing, hypertext glossing, etc. are used in EFL/ESL materials. Hulstijn and Laufer (2001a) are among the researchers who acknowledge the importance of glossing; however they also argue that although glossing can result in better IVA as opposed to no vocabulary aid conditions, it is less effective than dictionary consultation on the grounds that glossing requires less involvement load (i.e. the amount of mental effort needed to complete a task) than dictionary consultation does. Hulstijn and Laufer (2001a) operationalize "involvement load" with three cognitive components: (1) need, (2) search and (3) evaluation. They highlight that learners feel the *need* to know the unknown words and therefore *search* for their meanings while reading L2 texts with dictionaries at their disposal whereas they are spoon-fed when they are provided with glosses. They claim that dictionary use, which includes both the search and the need factors, must result in better vocabulary learning than glossed texts, which usually have the need factor only. As widely discussed as their theory is in the literature, there is no consensus on the comparative efficacy of glosses and dictionary use in IVA due to the limited amount of empirical evidence and the inconsistency in the findings of the available studies.

1.1. Statement of the Research Problem

While there are numerous studies on the efficacy of glosses (e.g. Cheng, 2009; Farvardin & Biria, 2012; Jacobs, Dufon, & Hong, 1994; Ko, 2012; Xu, 2010a) and on that of dictionary use (e.g. Chen, 2010, 2011; Cho & Krashen, 1994; Dziemianko, 2010; Knight,

1994; Lew, 2011; Luppescu & Day, 1993), the number of studies comparing the two conditions is very limited.

Interestingly, the few studies comparing glosses and dictionary use found contradictory results. Hulstijn, Hollander, and Graidanus (1996) gave their participants a reading passage (either with marginal glosses or a paper dictionary) and informed them that there would be a comprehension test on it but gave them a vocabulary test instead. They found that the marginal gloss group outperformed the paper dictionary group in the L2 into L1 word translation test. However, they also noted that the dictionary group performed far better than the other group with the words that they actually looked up in the dictionary. This problem, which is also pointed out by Hulstijn et al., indicates that the participants did not feel the 'need' to look up the words for the completion of the task. Xu (2010a) replicated the study carried out by Hulstijn et al. with only one major change. She added a delayed word translation (L2 into L1) test to her research design. The findings of her study were very similar to those of Hulstijn et al. in that she found the gloss group outperformed the dictionary group. However, she saw that the difference in the performances of the two groups disappeared in the delayed test. In her second study addressing the same question, Xu (2010b) compared the use of electronic dictionaries (as opposed to paper dictionaries which might have discouraged the participants in her previous study) to that of marginal glosses and found that the electronic dictionary group performed better than the gloss group in both the immediate and the delayed posttests.

The differences in the results of the aforementioned studies can be attributed to the intervening factors in each study. Firstly, none of the previous studies controlled for the need factor, which is seen as an indispensable part of IVA by Hulstijn and Laufer (2001a). In other words, whether or not the dictionary group consulted their dictionaries was not due to the requirements of the tasks but was rather the participants' level of motivation. Secondly,

considering the role that student motivation plays, the use of paper dictionaries could have discouraged the participants and influenced the findings of the studies. The one study that used electronic dictionaries (Xu, 2010a) had no control group, which means that the factor of lexical inference could have played an extraneous role as the inferability of the target words was not controlled for by the researcher. All this indicates the need for further research that will control for the factors that could have played an extraneous role in previous studies.

The present study will not only control for these extraneous variables but also expand the scope of previous research by measuring the participants' learning gains in different aspects of the target words. All of the aforementioned studies on glossing and dictionary use evaluated the participants' performances through tests that measured their receptive knowledge of meaning only. Researchers, however, agree that 'knowing a word' requires a lot more than mere knowledge of the meaning (Aitchison, 1994; Laufer, 1998; Miller, 1999; Nation, 1990, 2001; Webb, 2005, 2007). Indeed, Nation (2001) lists nine different aspects of vocabulary knowledge, which are (1) form and meaning, (2) concept and referents, (3) associations, (4) spoken form, (5) written form, (6) word parts, (7) grammatical functions, (8) collocation, and (9) constraints on use. Although studies exploring glosses and dictionary use have not looked into any of Nation's aspects except for receptive knowledge of meaning, there are many studies measuring the incidental gains of these aspects through other tasks (i.e. sentence writing, cloze tests, translation tasks, etc.). Among these studies, Webb (2007), focusing on the effect of frequency on IVA, stands out with 10 different posttests in its design. Webb looked at five different aspects of the vocabulary gains of the participants (i.e. orthography, association, grammatical functions, syntax, and meaning and form) on both receptive and productive levels. The present study takes Webb's design as a base but will look at only two of these five features, which are *meaning* and *word class* on two levels: (1) *receptive knowledge* and (2) *productive knowledge*. The reason why these two are chosen and

the rest of the features are not explored in this study is that the participants were exposed to the target words only once unlike the participants in Webb's study, who encountered the target words between 0 to 10 times. Webb noted that participants who encountered the words once performed well in meaning and word class tests while they could not perform well on the tests that required more complex gains such as association. Therefore, for the present study, only two most basic features that were likely to be acquired by the participants after one encounter were measured.

Lastly, this study compares specific types of glossing and dictionary use: L2 marginal glosses (i.e. vocabulary help in the margins) and online monolingual dictionaries - the two most common vocabulary aid forms utilized in the ESL context where the study is situated.

1.2. Research Questions

This study addresses the following research questions:

Are there any effects of L2 marginal glosses vs. online monolingual dictionaries used during reading comprehension tasks on ESL learners' immediate and delayed incidental acquisition of

1. receptive knowledge of meaning and word class?
2. productive knowledge of meaning and word class?

1.3. Importance of the Study

Reading and vocabulary tasks go hand in hand in many ESL/EFL classes. Even though lexical inference skills help learners to a certain extent, they still need to consult external aids for the meanings of some words. Some texts have glosses and some others do

not, in which case students have to use dictionaries. Therefore glossing and dictionary consultation exist in an either-or scenario. The present study has findings as to which of the two techniques are more helpful to learners. The results of the study will hopefully give some ideas to teachers who might want to modify their reading materials in a way that promotes IVA.

1.4. Delimitations

The present study has findings regarding the efficacy of marginal glosses and online dictionary use as two common techniques used in reading classes and it does not investigate how effectively these two techniques are used by students with different levels of motivation. Also, it does not look at students' dictionary use strategies or preferences. Lastly, the study explores students' knowledge of the meaning and word class of the target words only and does not offer any findings about the acquisition of other features of the target words.

1.5. Definitions of Constructs

Below are short definitions of some of the important constructs in this study. More detailed definitions, as well as how they are operationalized, are included in the literature review.

Incidental Vocabulary Acquisition: Hulstijn (1990) defines incidental vocabulary learning as a notion that "cannot be described in positive terms; it can only be described as the absence of its antonym i.e. the absence of an intention to commit something to memory and not to forget it'' (p. 33). In general terms, incidental vocabulary learning refers to the learning of vocabulary as a by-product of a task that is not aimed to teach lexical knowledge

(Krashen, 1989). In other words, learners acquire new words during a task that they do without having the intention of lexical learning (Bruton, Lopez, & Beta, 2011; Hulstijn, 1990; Hulstijn & Laufer, 2001a, 2001b).

Receptive Vocabulary Knowledge of Meaning (Rmeaning): Rmeaning is defined as the information a learner needs to know about a word in order to be able to understand its function and meaning when encountered in receptive tasks such as listening and reading (Crow, 1986; Mondria & Wiersma, 2004; Zhou, 2010).

Productive Vocabulary Knowledge of Meaning (Pmeaning): Pmeaning is defined as the information a learner needs to know about a word's meaning in order to be able to use it in productive tasks such as writing and speaking (Crow, 1986; Mondria & Wiersma, 2004; Zhou, 2010). In this study, the productive knowledge of meaning test required the participants to recall the words themselves when given a definition/synonym, which is a concept previously adopted by Webb (2007). As the participants encountered the words only once in the during-test, the tests asked for the most basic productive knowledge of the words, which may be considered close to the receptive side of the vocabulary knowledge continuum. (see. *2.3.2.4. Receptive and Productive Vocabulary Knowledge* for the receptive-productive knowledge continuum)

Receptive Vocabulary Knowledge of Word Class (Rclass): Rclass refers to participants' ability to recall the word class (i.e. verb, noun, etc.) of the target words when encountered.

Productive Vocabulary Knowledge of Word Class (Pclass): Pclass refers to participants' ability to use the target words in the correct syntactic role in productive tasks. This test, which was adopted from Webb's design (2007), measures very basic productive knowledge of word class and does not test the knowledge of association.

1.6. Definitions of Variables

The following variables are found in the design of the present study.

Independent Variable:

Type of Vocabulary Aid during Reading

- The Use of Marginal L2 glosses: One group was given a passage with eight target words, which were underlined and defined in the margin. One or two-word definitions of each word were given.

- The Use of Monolingual Online Dictionaries: One group was given the same reading passage with no modifications. They had access to Oxford American English Dictionary Online.

Dependent Variables:

- Receptive Knowledge of Meaning: The participants were expected to recognize the target words when encountered and recall their meanings (either in L1 or L2).

- Productive Knowledge of Meaning: The participants were expected to recall the target words themselves when they are given the L2 definition.

- Receptive Knowledge of Word Class: The participants were expected to make judgments about the syntactic accuracy of given sentences.

- Productive Knowledge of Word Class: The participants were expected to write sentences by using the target words in correct syntactic roles.

Control Variables:

- The factor of need: The reading comprehension questions required the participants to look up the target words and thus the factor of need was controlled for. A special version of the passage was prepared and the target

words were replaced with nonsense words. A native speaker and an Egyptian proficient speaker of English were asked to read the text and answer the comprehension questions that required the knowledge of the target words (the nonsense words, in their case). They both reported that they could not guess the meanings of the words from context and therefore could not answer the questions. This was taken as an indication that the comprehension questions could not be answered without the knowledge of the target words.

- Lexical Inference: The target words were presented in the passage in a way that did not allow the participants to infer their meanings from the context. The two colleagues who took the special version of the test reported that the passage did not lend itself to lexical inference for the target words. One of these two colleagues was a proficient Egyptian speaker of English and that she could not infer the meanings of the words was taken as an indication that the participants would not be able to infer the meanings of the target words.

- Morphological clues: The target words did not have any morphological clues so the participants had to rely on the information gained from the glosses and online dictionaries for the completion of the posttest. None of the target words had any affixes (i.e. un-, in-, -able, -ly, etc.), which might have helped them with the two tests on word class.

- Topic Familiarity: A passage on an unfamiliar topic (Atlantis) was intentionally chosen in order to eliminate the unwanted effect of previous knowledge about the topic, which could have helped the participants guess the meanings of the target words. A questionnaire item asking whether participants were familiar with the topic before they read the text or not was included at the end of the comprehension questions. The answers showed that none of the participants were familiar with the topic, except for two, whose

scores were not counted.

The study did not have a control group for three reasons. Firstly, the above-mentioned control variables suggest that the most important of the factors that might have affected the participants' vocabulary acquisition were already controlled for, which made having a control group unnecessary. Secondly, the pretest of this study makes one confident that the results were due to the treatment only. Lastly the aim of the study was to look at the comparative effects of dictionary use and marginal glosses, whose individual positive effects on IVA have already been shown in many previous studies. Therefore, the study had two experimental groups whose scores were compared to each other's.

1.7. Terms and Abbreviations

Considering the frequency of the use of some constructs and terms, the following abbreviations will be used in the thesis:

IVA – Incidental vocabulary acquisition

L2 – Second language (English)

L1 – Native language (Arabic)

Rmeaning – Receptive knowledge of meaning

Pmeaning – Productive knowledge of meaning

Rclass – Receptive knowledge of word class

Pclass – Productive knowledge of word class

Chapter II

2. Literature Review

2.1. Introduction to the Literature Review

This literature review concentrates on the pedagogical and theoretical foundations of the use of glosses and dictionaries in IVA through reading tasks. It is intended to (1) offer an exploration of various definitions of some important concepts in IVA, (2) explain the perspective the present study adapts, and (3) provide examples from the studies carried out on related topics.

2.2. Nature and Structure of the Literature Review

The present literature review is divided into three sections. The first part is dedicated to the pedagogical necessity of comparing the use of glosses to that of dictionaries and explains the frequency of the use of each vocabulary aid. The second part concentrates on the underlying theories and recent discussions in the literature while the last part gives an overview of studies that looked into the use of glosses and dictionaries. The findings of the studies mentioned in the literature review are organized thematically rather than chronologically. However the importance of the sequence of the studies is highlighted when necessary.

2.3. Comparing Vocabulary Glosses to Dictionary Use

2.3.1. Pedagogical foundations

Research shows that a great amount of L2 vocabulary is acquired incidentally during different reading tasks and extensive reading (Keating, 2008; Krashen, 1989; Read, 2000). It is argued that learners need to know around 98% of the words in a text in order to be able to understand a text successfully (Hu & Nation, 2000; Nation, 2001). When learners encounter an unknown word that they deem important for their comprehension of the text, they either (1) ignore the word, (2) consult an external aid (i.e. teacher, friend, dictionary, gloss) or (3) try to infer the meaning of the unknown word from the word's immediate environment (Laufer & Yano, 2001). In his taxonomy of L2 vocabulary learning strategies, Schmitt (1997) noted that lexical inference, or what he calls "discovery", was the most common strategy used by ESL learners. Guessing, or lexical inference, which is widely used by language learners and promoted by language teachers, is indicated to be the first strategy that a learner uses upon encountering an unknown word unless s/he ignores it completely (Paribakht & Wesche, 1999; Qian, 2004). Research has shown that when learners try to infer the meanings of unknown words, they get engaged and they think deeply to reach the meaning, which results in better acquisition of the unknown words (Hulstijn, et al., 1996; Watanabe, 1997)

Even though lexical inference is shown to be a highly effective technique, research indicates that learners need to meet certain preconditions for successful lexical inference. Firstly, a good level of vocabulary knowledge is crucial for lexical inference. Lui and Nation (1985) argue that learners need to know about 95% of the words in a text in order to correctly guess the meanings of the rest. The literature also shows that learner proficiency, as well as mental lexicon size, is a significant variable in learner's ability to infer and that inference is most likely to be accurate when used by high proficiency learners (Paribakht, 2005). Thirdly,

learners need to have certain knowledge about word forms. Nation (2006) states that in order to be able to correctly infer the meaning of a word, learners must be familiar with 8000-9000 word families. Fourthly, lexical inference requires a developed ability to think critically, which Husltijn (1993) argues is not owned by some learners. Lastly, it has been recently suggested that cultural and contextual familiarity or background knowledge plays a decisive role in learners' ability to infer meaning (Pulido, 2004; 2007). The findings of Pulido's two studies on this issue (2004; 2007) show that learners are more accurate in their lexical guessing while reading passages on familiar topics.

Even though accurate lexical inference is shown to result in better vocabulary acquisition, its preconditions are not always met by learners (Hulstijn, 1993; Mondria, 2003). Not all learners have the same background knowledge or inference ability and not all words lend themselves to correct inference (Laufer, 1997). Therefore, when all the other options fail (i.e. lexical inference, asking a teacher or a peer), learners are left with only two options: (1) referring to glosses when available or (2) consulting a dictionary. Indeed, even when they rely on their inference, learners are recommended to look up the words in order to be sure of its meaning not only because wrong inferences are common (Hulstijn, 1992; Hustijn et al., 1996) but also because learners usually overrate their ability to infer meanings even when most of their inferences are wrong (Kaivanpanah & Alavi, 2008; Laufer & Yano, 2001).

Thus, dictionary consultation and glosses are left as two strategies for learners especially when there is not a teacher or a peer present. While glosses are widely present in various forms in ESL/EFL materials today (Ko, 2012), the use of various types of dictionaries still preserves its importance among language learners (Chen, 2011). Comparing these two common techniques is highly crucial in that such a comparison is likely to offer information about which one of these two techniques is more effective when incidental vocabulary acquisition is aimed. As a result of such a comparison, the more effective of the two

techniques can be promoted by teachers and material designers so that learners are provided with a more fruitful vocabulary acquisition environment during reading tasks.

2.3.2. Theoretical foundations

The present study is based on many important concepts and theories. Before moving onto the methodology section, it is crucial to explain how these concepts are defined and how they relate to the present study. The following section explains the role of glossing and dictionary use in different important ESL/EFL theories and shows the theoretical perspectives that are adapted in this study.

2.3.2.1. Incidental and intentional vocabulary acquisition

Even though the term "incidental" is very frequently encountered in the field of L2 vocabulary acquisition, there is no consensus as to how the term can be defined. According to the most common of the different definitions available, IVA is described as acquiring new words as a by-product of another task (usually reading and listening). As two important supporters of this definition, Huckin and Coady (1999, p. 182) clearly state, incidental acquisition is "a by-product, not the target, of the main cognitive activity". This definition is adapted by many researchers and many studies were carried out in the field of IVA based on this definition (e.g. Day, Omura, & Hiramatsu, 1991; Laufer & Hill, 2000; Mondria, 2003; Read, 2000; Paribakht & Wesche, 1999; Wode, 1999). However, it is argued that defining IVA as a "by-product" creates rather a vague image and does not provide any explanation regarding what actually decides whether a task is considered as the main goal or as a by-product (Gass, 1999; Bruton, Lopez, & Mesa, 2011).

Hustijn (1990) acknowledges this lack of clarity in the definition of IVA and states that IVA is a notion that "cannot be described in positive terms; it can only be described as the absence of its antonym, i.e. the absence of an intention to commit something to memory and not to forget it'' (p.33). Similarly, Schmitt argues that IVA is learning a form when there is not a primary objective to learn it (1997). The idea of considering IVA as a result of a lack of intention to learn new words is supported by many researchers (i.e. Barcroft, 2004; Hustijn, Hollander, & Greidanus, 1996). However, it is hard to say that this definition is found sufficient, either.

Gass (1999) and Ellis (1994) called attention to that the term "incidental" is directly related to the cognitive processing of words and they conclude that whether learning is incidental or not cannot be controlled in any ways. Gass (1999, p.320) says "There is no way to show that a given word was incidentally learned (…) (because) it is not possible to have direct access to what a learner is doing." Ellis (1994, p.7) also highlights the ambiguity of the term and points out "that we have not been taught vocabulary does not entail that we have not taught ourselves". Therefore, incidental acquisition is seen as rather an abstract term that is hard to operationalize due to learners' individual differences and learning strategies that are beyond the access of researchers.

The only operational definition for IVA is suggested by Hulstijn (2006), who describes incidental and intentional vocabulary acquisition in this way:

> Intentional learning refers to the learning mode in which participants are informed, prior to their engagement in a learning task, that they will be tested afterward on their retention of a particular type of information. Incidental learning refers to the mode in which participants are not forewarned of an upcoming retention test for a particular type of information. (Hulstijn, as cited in Alemi & Tayebi, 2011, p. 83)

This definition only applies to classroom context as opposed to the other definitions mentioned above. While researchers agree that most vocabulary is acquired outside class and it is obviously harder to decide whether learning is incidental or not as learners are left alone with the material outside class, classroom content learning is more advantageous for research purposes as learners are more closely observed by researchers. Nevertheless, the fact that learning takes place in class does not guarantee that a researcher can control the type of learning as there is always a possibility that learners might use individual vocabulary learning strategies even when they are not asked to focus on it.

Nonetheless, what can be concluded from the available literature is that although it is mostly agreed that incidental learning might not be totally incidental for some learners who might or might not utilize individual vocabulary learning strategies during different tasks, whether a task itself requires the learner to concentrate on vocabulary for an imminent test or not is seen as what decides whether learning is incidental or intentional.

Given the elusiveness of the two terms, many researchers prefer to see incidental and intentional learning as two sides of a continuum rather than two constructs existing in a dichotomy (Gass, 1999). The present study adapts Hulstijn's abovementioned operational definition of incidental learning and was designed accordingly, like many other studies in the field (Pichette, De Serres, & LaFontaine, 2011; Webb, 2005, 2007; Zandieh, 2012)

2.3.2.2. Noticing and attention

Schmidt (1993) argues that whatever is acquired - intentionally or incidentally- is what is noticed. Even though the term "incidental" might mean unintended, it does not mean "unattended" (Laufer & Hill, 2000, p. 58). Gass (1999, p.321) highlights that incidental acquisition "does not mean that the learner does not notice the word in question." Therefore,

IVA should not be seen as learning without noticing. At this point, it is necessary to look at how noticing and IVA can exist in the same context and in what ways noticing affects IVA.

According to Laufer and Hill (2000), in order for IVA to be effective, the learner should go through a conscious learning process. Accordingly, while vocabulary acquisition should not be the learner's main aim, necessary aid in vocabulary should be provided to the learner so that IVA is triggered. Laufer and Hill (2000) state that looking up a word or referring to glosses, as well as inferring its meaning, do not make learning any less incidental but rather promotes better incidental learning.

In addition to different types of lexical help provided to learners, the relevance of the vocabulary items to the main task is seen as a successful way of integrating noticing into IVA. In his 1993 study, Hulstijn shows that learners acquire words that are made noticeable to them through the requirements of reading tasks more effectively. He notices that words that are relevant to the reading task (therefore more noticeable) are better acquired than those that are not. He also states that the nature of the task remains incidental as learners noticed the words not for an upcoming vocabulary test but rather for the completion of the main task.

Hence, the present study sees noticing as a necessary component of IVA and takes the relevance of the target words to the task and providing vocabulary aid as two factors that are crucial to lead learners to notice the new forms, which does not make the learning process an intentional one.

2.3.2.3. Task induced involvement and learning burden

While the definition of the term "incidental" remains to be a controversial one in the literature (Bruton, et al., 2011), the number of studies in the field of IVA increases every day.

Therefore, the "lack of intention" that is recurrent in various definitions of IVA only refers to the lack of intention for learning new words on the part of learners. For teachers and researchers, there is nearly always a quest as to how an ideal environment for IVA can be provided to learners. This investigation regarding the effectiveness of incidental vocabulary learning entails a more specific question: what kind of tasks result in better incidental vocabulary acquisition?

The input hypothesis that was proposed in 1982 by Krashen had a considerable effect on the understanding of vocabulary learning process (Xu, 2010a). Krashen (1985) suggests that learners acquire L2 vocabulary incidentally when they are exposed to L2 materials. Accordingly, reading is seen as an effective tool for L2 IVA. As opposed to the input hypothesis, the output hypothesis puts emphasis on the importance of learners' production in L2 (Swain, 1985). According to the output hypothesis, language acquisition can occur most successfully if learners focus on productive tasks such as writing (Ellis & He, 1999; Swain, 1985).

While there are numerous studies comparing input and output tasks for their relative efficacy in IVA, one study that offers an alternative answer to the question was carried out by Hulstijn and Laufer (2001a). The hypothesis they propose, the Involvement Load Hypothesis, is most commonly investigated and discussed by researchers (eg. Hulstijn & Laufer 2001b; Keating, 2008; Pichette, et al., 2012). Hulstijn and Laufer (2001a) argue that neither the input nor the output hypotheses can explain L2 vocabulary learning process. Instead, they investigate what kinds of factors play a role in a learner's performance in vocabulary learning and they decide that "involvement load" as a construct is actually what results in better L2 learning.

What Hulstijn and Laufer call "task induced involvement" or "involvement load" is a reminiscent of Craik and Lockhart's levels of processing theory (1972), which is based on the idea that the degree to which new forms are remembered by the learner is decided by how deeply he/she analyzes them. Similarly, Schmitt and Schmitt (1995) argue that words that are acquired with less effort are forgotten more easily.

Nation (1990; 2001) takes these ideas one step further and concentrates on the features of the target items rather than the learner. He argues that each word has a *learning burden*, which decides how much effort the learner needs to exert to learn it. Nation (1990; 2001) points out that the learning burden each word has is decided by the learner's cultural familiarity, background, and previous knowledge of the word class and morphological cues.

One important factor that makes Hulstijn and Laufer's involvement load theory different from the abovementioned ones is that they operationalize their theory with three cognitive components: *need*, *search* and *evaluation*. The need factor can be explained as learners' need to know the meaning of a certain L2 word. The search factor is the process through which learners reach the meaning of a target word. The more they search for the meaning, the more involved they get and therefore the better they learn (Hulstijn & Laufer, 2001a). Lastly, learners should also evaluate L2 words; that is to say, they need to compare the target words to some other words they know and, therefore they should gain knowledge about how, when and where to use the new words. According to Hulstijn and Laufer (2001a), if a task requires all these three components, it means that the task encourages more involvement load than those that require only one or two of them. Likewise if a task requires only two of these three components, it results in better IVA than those that require only one.

Hulstijn and Laufer (2001a) also offer a list of different tasks through which incidental vocabulary learning is accomplished. They measure the involvement load that each

task requires (i.e. need, search, and evaluation) and rank them according to their expected effectiveness in vocabulary learning. In other words, they offer a hierarchy of tasks in terms of their efficacy in incidental vocabulary acquisition. The proposed hierarchy has been investigated many times by various researchers regarding different aspects of vocabulary acquisition. Some researchers compared the efficacy of writing and reading tasks (Pichette, et al., 2012; Webb, 2005); some compared the effects of different writing tasks (Folse, 2006; Liu, 2008); some others investigated the effect of an integration of reading and writing tasks (eg. Hulstijn & Laufer, 2001b; Keating, 2008; Pichette, et al., 2012); and still others explored the effects of different reading conditions (Ko, 2012; Webb 2007; Xu, 2009; 2010b), which is the focus of the present study.

According to Hulstijn and Laufer (2001a), reading L2 texts with glosses requires less involvement load than reading them with a dictionary at one's disposal as learners feel the need to know the target words and therefore search for their meanings in the latter while they are spoon-fed in the former. They argued that dictionary use, which includes both the search and the need factor, must result in better vocabulary learning than glossed texts, which have the need factor only.

This leads the reader to question whether glossing unfamiliar words in reading texts, which is a common practice in TESL/TEFL, and looking up unknown words in a dictionary are effective tools when incidental vocabulary acquisition is aimed at by the teacher. The studies looking into this comparison will be discussed in the following sections (see *Studies on Glossing and Dictionary Use*)

2.3.2.4. Receptive and productive vocabulary knowledge

The distinction between receptive and productive vocabulary knowledge is usually seen as a reflection of the difference between reading/listening tasks and writing/speaking tasks (Crow, 1986; Nation, 2001). Receptive vocabulary knowledge reflects to what extent learners can recognize the features of a word when they encounter it in reading or listening tasks while productive knowledge shows to what extent they can use the word in writing and speaking.

Meara (1990) defines receptive and productive vocabulary knowledge as passive and active knowledge and argues that passive knowledge (receptive) can only be activated by external stimuli (the written or oral form of the word) whereas active knowledge is activated through some associations that one word has with other words in the learner's mental lexicon (i.e. L1 equivalent or related words).

Stoddard (1929) is one of the first researchers who looked into the difference between the two types of vocabulary knowledge and found some important result which are still discussed in the literature. After comparing his participants' receptive and productive vocabulary knowledge through L1 into L2 and L2 into L1 translation tests, Stoddard found that the receptive vocabulary sizes of the participants were always larger than their productive vocabulary sizes. Other studies conducted about the same question revealed similar results (e.g. Fan, 2000; Laufer, 1998; Webb, 2008). Nation (2001) points out that receptive vocabulary knowledge is always larger than productive knowledge indicates that the two types of knowledge should be seen as two sides of a continuum. It is argued that learners nearly always start with the receptive side of the continuum and move towards the productive side as they improve their proficiency. The role of reading in helping learners move from the receptive side to the productive side is discussed in the following part.

2.3.2.5. Different vocabulary features acquired during reading

In relation to the discussion about receptive and productive knowledge, Read (2000) questions the operationalization of the two constructs and asks where a possible threshold can be placed between the two types of vocabulary knowledge and how these constructs can be tested. One of the first studies looking at different testing methods of vocabulary was carried out by Richards (1976). He argues that there are seven different types of knowledge a learner has with regard to each word and states that each should be tested and these are knowledge of (1) association, (2) limitations, (3) syntactic behavior, (4) possible derivations, (5) main meaning, (6) other meanings, (7) frequency of encounter. Nation (1990) offers a similar but more detailed list for vocabulary knowledge features and lists 16 categories on which learners should be tested. The list was shortened and made more concise by Nation (2001) after a decade and vocabulary knowledge was divided into nine aspects (i.e. spoken form, written form, word parts, connecting form and meaning, concepts and referents, associations, grammatical functions, collocations, constraints on use).

Most of the research on the acquisition of different features of vocabulary is carried out in the field of extensive reading. In their 2006 study, Pigada and Schmitt measured the vocabulary gains of their participants after one month of reading. The authors tested the participants' knowledge of the meaning, spelling and grammatical functions of the target words and reported that there were significant learning gains in all three types of word knowledge with spelling being the earliest acquired one.

Pellicer-Sanchez and Schmitt (2010) carried out a similar study and explored the IVA of ten target words in an authentic novel, *Things Fall Apart* by Achebe. After the reading stage was completed, the participants were given four main posttests: spelling recognition, word class recognition, meaning recall and meaning recognition. The results showed that the

participants had considerable learning gains in all four aspects measured. The researcher reported that meaning recognition was the best acquired one, which was followed by spelling recognition, word class recall and meaning recall respectively.

Webb (2007) tested ten different word features acquired in relation to the number of encounters and he conducted ten posttests (i.e. spelling, meaning, syntax, association and grammatical functions both on receptive and productive levels). He found that the more times the participants encountered the target words, the better they performed on all tests. In compliance with all the other mentioned studies on the same topic, Webb noted that spelling, word class and meaning were the easiest acquired features of all.

While the abovementioned studies differ from many other studies in the field of IVA through reading in that they measure the acquisition of different word features in their posttests, they did not look at the role of vocabulary aid (dictionaries, glosses, teacher or peer help, etc.) in the incidental vocabulary gains of the participants. Measuring different word features is more common in the research on dictionary use than in the research on glossing. The few studies that looked into this area were limited to the productive knowledge of meaning and collocations acquired with the help of dictionaries and did not go any further (e.g. Dziemianko, 2010; Laufer, 2011). The studies on glossing, on the other hand, were even more limited in this sense and always looked into the receptive gains of meaning by learners (Ko, 2012). There are, to my knowledge knowledge, no studies looking at various word features acquired through dictionary use vs. glossing.

2.4. Studies on Glosses and Dictionary Use

The following parts investigate the results of various studies carried out on the effects of glosses and dictionary use on IVA.

2.4.1. Studies on glossing

As a common example of modified input, glossing is used very frequently in ESL/EFL classrooms and has been a focus of research for the last few decades. (Ko, 2012). Nation (2001) describes glossing as providing short descriptions or synonyms –L1 or L2 – of important words in a reading passage. Although the definition of glossing seems to be clear enough, it is possible to find various forms of glosses today - such as hyper glosses, multiple choice glosses, picture glosses – drawing a far bigger picture than the one that is offered in Nation's definition.

Research has shown a facilitating effect for these various types of glossing on reading comprehension, and in decreasing the time spent on reading tasks (Ko, 2005). While many studies concentrate on the relationship between glossing and reading comprehension, glossing is of a fundamental role in IVA, as well. Studies comparing glossed texts to those with no modifications found a facilitating effect of glossing on vocabulary acquisition (Cheng, 2009; Jacobs, Dufon & Hong, 1994; Farvardin & Biria, 2012; Ko, 2012; Xu, 2010a). This consensus led to a deeper analysis of glosses and many researchers have looked into factors affecting the already-acknowledged efficacy of glosses, such as the importance of the place of glosses in a reading passage, the language of glosses and the effect of enriched (MCQ and picture) glosses. The findings of these studies are briefly explained below.

The first important factor that has been investigated by researchers is the language of glosses. Jacobs, Dufon, and Hong (1994) are among the first researchers exploring possible differences between L1 and L2 glosses. The participants of their study were 85 native speakers of English who were learning Spanish as a second language at a university. The participants were randomly assigned to one of the three groups of the experiment. They were asked to read a 623-word passage either with L1 glosses (English) or L2 glosses (Spanish) or

no glosses. No statistically significant difference was seen between the immediate and delayed results of L1 and L2 groups. Although the results of the tests did not show any superiority of L2 glosses, most of the participants of the study preferred L2 glosses over L1 ones. However, many of them noted that L2 glosses were helpful only when they were comprehensible, which is directly related to learners' level of proficiency.

Xu (2010) compared the recall performance of the participants who were given (1) L1 glosses, (2) L2 glosses or (3) both L1 and L2 glosses during an L2 reading task. In the immediate test, words that were glossed in L1 were remembered more successfully than those that were glossed both in L1 and L2, which in turn had a better recall rate than those that were glossed only in L2. Xu interviewed some of the participants after the posttest and most of them mentioned that they usually ignored the L2 definition either because the L1 equivalent was clear enough or because it was shorter than the L2 definition.

More recently, Ko (2012) addressed the same issue in her study in which 90 Korean students studying English at a university in South Korea participated. They were asked to read an L2 text under one of the three conditions: (1) L1 glosses, (2) L2 glosses, or (3) no glosses. The findings of this study were not any different from those of Jacobs et al.'s. There was no statistically significant difference between the L1 and L2 groups either in the immediate or in the delayed test. However, just like what Jacobs et al. found in their study, participants said that they favored L2 glosses over L1 glosses in the questionnaire that they were given after the posttest even though the L2 group did not perform better than the L1 group. Clearly, the results of the studies on the language of glosses do not show a consistent superiority of one condition over the other while learners tend to be more in favor of L2 glosses than L1 ones.

In relation to the place of glosses, which is another important factor that has frequently been explored, research shows that vocabulary aid during reading leads to better results than before or after reading (Alessi & Dwyer, 2008). As an exception, Holly and King (1971) found no differences among marginal glosses, end-of-page glosses and wordlists attached to the text with regard to their effectiveness either in reading comprehension or IVA. Differently from the rest of the studies, Ko (2005) looked at learner preferences with regard to the place of glosses and noted that learners tend to favor marginal glosses over other forms even though she did not report any superiority of one type over the other.

While glossing is considered as a helpful tool for IVA as opposed to no gloss conditions, it is also argued that their efficacy can be increased if the content of glosses is enriched. Hulstijn (1992) argues that glossing does not require as much mental effort as lexical inference does and suggests the use of multiple choice glosses to increase the work load and therefore encourage deeper processing. In MCQ glosses, learners are given at least one distracter word along with the correct definition and they are asked to consider both options and choose the one that fits the context.

Watanabe (1997) compared single glosses to MCQ glosses and no gloss condition and reported no statistically significant differences between the performances of the single gloss and the MCQ gloss groups whereas the two groups performed better than the no gloss group. Nagata (1999) investigated the same question in a CALL context and compared single and MCQ glosses in online reading tasks. As opposed to Watanabe's findings (1997) and supporting Hulstijn's argument (1992), she found a significant superiority of MCQ online glosses over single online glosses in IVA. More recently, Farvardin and Biria (2012) measured the effects of enhanced glosses and they divided the participants, who were 120 Persian speaking EFL learners, into three groups and had them read the same text under three different conditions: single L1 glosses, single L2 glosses, and multiple-choice L2 glosses.

The participants were given the correct L2 synonyms with a distracter under the third condition. Neither of the two single gloss groups did as well as the multiple-choice L2 group, which supports Hulstijn and Nagata.

Looking for other possible ways of enriching glosses, Yoshii (2006) had his participants, who were 192 Japanese speaking EFL learners, read an English text under one of the four conditions: (1) L1 glosses only; (2) L2 glosses only; (3) L1 glosses plus pictures; and (4) L2 glosses plus pictures. Immediate and delayed MCQs and translation tests were administered. Yoshii saw that both L1 and L2 glosses were far more helpful in vocabulary acquisition when they were combined with pictures. Similarly, Shahrokni (2009) had three experimental groups who were given (1) text glosses, (2) picture glosses, and (3) text and picture glosses and reported that the last condition; that is the most enriched one, resulted in better results than the other two.

What was common in all these studies was that (1) they all proved glossing, no matter what type, to be helpful in IVA compared to no-gloss conditions and that (2) they showed the efficacy of glosses could be increased through designing conditions that require more effort on the part of the learner.

2.4.2. Studies on dictionary use

Investigating the involvement load hypothesis, Keating (2008) argues that reading with a dictionary requires more involvement load than any other vocabulary aid learners might use as looking up a word in a dictionary requires learners to notice the word, temporarily isolate it from the context, find the corresponding entry in the dictionary, read all possible meanings, decide on the one that is most likely to occur in the given context and go

back to the main text to make sure that the meaning fits the context. Keating (2008) states that dictionary use requires all the three components of involvement load (i.e. need, search, and evaluation) and therefore must have a significantly positive effect on vocabulary acquisition. Crookall and Oxford (1990) propose a similar argument about the positive effect of isolating words for looking them up on increasing the focus on form for IVA.

In an attempt to investigate the reliability of these theories about the role of dictionary use in IVA, Luppescu and Day (1993) carried out a quantitative study that they argue to be the first one offering empirical evidence about the dictionary use and IVA relationship. Their participants were 293 first and second year college students in Japan, who were all studying English. This rather large sample was divided into two groups. The treatment group was asked to read a short story with a bilingual dictionary at their disposal and the control group relied on their lexical inferring skills. The short story included 16 target words that were reported to be unknown by the participants in a pretest of recognition. No time limit was given to the participants. In the immediate surprise posttest, they were given multiple-choice L2 into L1 translation questions. The results of the posttest showed that the dictionary group performed better than the control group overall. However, Luppescu and Day (1993) noted that the performances of the two groups revealed inconsistency when the results for individual target items were investigated. They attribute this inconsistency to the possibility that some students might not have looked up some of the target words or that some target word entries in the dictionaries might have been misleading. They also mention that the latter is more likely as the students were allowed to choose any bilingual paper dictionary they wanted.

One year later, Knight (1994) addressed the same question and carried out a very influential study in which she divided 112 learners of Spanish into two experimental groups. The first group was asked to read four authentic Spanish articles with an online dictionary at

their disposal while the second group was given the same task with no dictionaries. The participants were informed that the task was done for measuring their reading comprehension skills and they were asked to read the articles and summarize what they remembered in English (L1). Just like Luppescu and Day's, Knight's study showed the superiority of the dictionary group over the lexical inference group both in the immediate and the delayed posttests of L2 into L1 translation. These findings have been supported by many other researchers who looked at the efficacy of using dictionaries (i.e. Cho & Krashen, 1994; Fraser, 1998; Lew, 2011).

Relying on the idea that dictionary use helps learners in IVA, some researchers compared different kinds of dictionaries in terms of their efficacy in vocabulary acquisition and investigated possible ways in which dictionaries can be made more useful to learners. The language of dictionary entries and the physical form of the dictionaries are two factors that have most commonly been explored by researchers and the results of these studies are summarized below.

Unlike the research on the language of glosses, the research on the language of dictionaries does not compare L1 and L2 entries but rather sees each option as a required method per se. Indeed, dictionary research in general is seen as a field that is most related to individual differences (Hartmann, as cited in Dziemianko, 2010) and therefore it is generally agreed that the more types of information provided in a dictionary, the more effective it is as learners have their own needs and expectations and that a dictionary with enriched entries has the advantage of meeting the needs of different learners during different tasks (Chen 2008; Laufer & Hill, 2000). In accordance with this idea, White (1997) argues that monolingual and bilingual dictionaries should not be seen in an either-or scenario but rather be considered as two options that can (and should) exist together and that the use of the latter can be seen as a considerable aid to the use of the former. Gouws (2004) supports this opinion by stating that

each dictionary type (monolingual or bilingual) serves to a different purpose in language learning and learners benefit from each at different times.

As an alternative to monolingual and bilingual dictionaries, bilingualized dictionaries offer the learner the ability to check both L1 and L2 definitions of the words and therefore are considered to be more user friendly than the rest. In their 1997 study, Laufer and Hadar compared their efficacy in increasing reading comprehension to that of bilingual and monolinagual dictionaries and saw that bilingualized dictionaries were far more useful to the participants.

Laufer and Levitzky-Aviad (2006) proposed the use of still another type of dictionary, which provided learners with L1 definitions, L2 definitions and L2 synonyms at the same time. They compared the effectiveness of (a) L1 →L2-L2 dictionaries, (b) L2 →L2 –L1 (bilingualized) dictionaries, and (c) L1→L2 (bilingual) dictionaries in IVA. 85 college students participated in the study and they were asked to translate 36 Hebrew (L1) sentences into English (L2) by using the aforementioned types of dictionaries. The study adapted a within-subjects design. The results showed a superiority of the L1 →L2-L2 dictionary group over the other two.   Laufer and Levitzky-Aviad's study reinforce the idea that the more information is provided the better for the learner as each learner needs and each task requires a different piece of information about the looked up words. More recently, Chen (2008) compared learner perspectives of monolingual, bilingual and bilingualized dictionaries and found that learners preferred each dictionary type for different purposes and that they deemed all types as necessary, indicating that learner preferences support what has already been suggested by researcher.

Just like the language of glosses, the form of dictionaries used by learners and their comparative effectiveness in IVA has been widely explored in the field of dictionary

research. As a natural result of the recent developments in CALL, paper dictionaries have been replaced by online-electronic dictionaries, pocket dictionaries and more recently with mobile phones that have dictionary functions. While there is an increasing amount of research on both pocket dictionaries (Jian, Sandnes, Law, Huang, & Huang, 2009) and mobile phone use (Lu, 2008), only the use of online-electronic dictionaries will be elaborated on here as it directly relates to the present study.

Aust, Kelley, and Roby (1993) compared conventional paper dictionaries to hypertext dictionary entries and found that hypertext electronic dictionaries were significantly more consulted by the participants than paper dictionaries and therefore led to better results. In their aforementioned study, Laufer and Levitzky-Aviad (2006) also compared paper and electronic L1 –L2 –L2 dictionaries and found that electronic versions of the entries were consulted more frequently than the paper versions even though there were no differences in their efficacy. Indeed, it has been claimed that the frequency of look-up and the user friendliness of electronic dictionaries might not always entail better IVA (Laufer and Hill, 2000).

In her 2010 study, Dziemianko called for attention to the lack of studies comparing the efficacy of online dictionaries in IVA and compared the paper and the online version of the Collins COBUILD dictionary. Sixty-four upper intermediate and advanced learners of English were divided into two treatment groups and were asked to complete two consecutive tasks. In the former task, they were given a list of English words and were required to supply a definition for each either in English or in Polish (L1) by using either the paper version or the online version of the dictionary. In the latter test, they were given preposition + noun collocations (i.e. *on the verge*) with the prepositions being deleted and were asked to write the correct preposition. The same tasks were given two weeks later for delayed results. Dziemienko reported that the electronic dictionary group performed significantly better than

the paper dictionary group in both tasks and the results preserved their significance in the delayed test. In a similar study, Chen (2011) achieved very similar findings. He compared paper and electronic bilingualized dictionaries and found that the latter had better retention of the target words than the former while there were no statistically significant differences found between the performances of the two groups in the immediate posttest. Clearly, electronic versions of dictionaries (monolingual, bilingual or bilingualized) encourage more look-ups and result in better IVA. They also require less time and therefore distract learners less than paper dictionaries do, which is pointed out as an important factor by Dziemienko. (2010)

2.4.3. Studies comparing glosses and dictionary use

Even though there is a great amount of research both on glossing and dictionary use – some of which has already been summarized above -, none of the studies mentioned in the two preceding sections compared glossed conditions to those with dictionaries In other words, none of them compared the effects of the different amounts of involvement load that glosses and dictionary use might have, which resulted in the lack of ample empirical evidence for the superiority of one condition over the other. The few studies that addressed this gap in the literature will be discusses in the following section with their findings and limitations.

Hulstijn, Hollander, and Greidanus (1996) are among the first researchers who compared the efficacy of glossing to that of dictionary use. They tested the effects of three conditions on the vocabulary gains of 78 Dutch advanced learners of French and these conditions were reading with (1) L1 marginal glosses, (2) access to bilingual dictionaries and (3) neither. All three groups were given a reading passage and were informed that they were going to be given a comprehension test afterwards, during which they would not be given access to the passage. When they finished, they were given a vocabulary translation test

(from L2 into L1) instead of a reading comprehension test. The findings of the study showed that the first group performed better in the post-test than the second and the third group.

Hulstijn et al. (1996) also looked at which words the participants looked up in the dictionary and they saw that they looked up only 12% of the target words in the passage, which suggests that the reading task was not successful in encouraging the participants to look up the words. When they limited the results of the study to those words that the dictionary group looked up, they found a higher retention rate in the dictionary group than in the rest, which shows that when new words are looked up in a dictionary, the likelihood of remembering them is higher.

Although these findings are quite significant, there are also some important limitations to be considered with regard to Hulstijn et al.'s (1996) study. First of all, the fact that they asked the participants to use paper dictionaries might have discouraged some from looking up the words as it necessitates more effort than online or electronic dictionaries. Secondly, they compared the use of L1 glosses to the use of bilingual dictionaries. Therefore the results of the study might have been affected by the difference in the language in which definitions of the words were offered. Thirdly, as the participants of the study were advanced learners, they might have used lexical inference rather than referring to the glosses or to the dictionaries. Lastly, even though Hulstijn et al. (1996) increased the factor of need in the reading task by informing the participants about an imminent comprehension test, the need for looking up the target words for the task might not have been clear enough to the students as they did not know what parts of the passage they were going to be tested on. Some of them might have been satisfied with what they understood from the passage without looking up the words.

Xu carried out two different studies addressing the same question as Hulstijn et al. (1996) did. In her first study, Xu (2010b) found a similar result to Hulstijn et al.'s (1996). 125 ESL learners were divided into four different groups. Xu (2010b) investigated the comparative effects on vocabulary acquisition of (1) reading with marginal glosses, (2) reading with marginal glosses plus making sentences with target words, (3) reading with a paper dictionary at one's disposal and (4) reading with no external aids. The participants were not given or informed about any comprehension tests in this study, which totally eliminated the factor of need. After they finished reading, they were given a surprise vocabulary translation task (from L2 into L1). Xu (2010b) found that the group who used marginal glosses performed better in the translation task than the dictionary group. She also pointed out that the difference between the scores of the two groups disappeared in the delayed test, which was given five days later than the immediate test.

In her second study, Xu (2010a) used a similar design that did not control the need factor; however she attempted to overcome the limitation seen in Hulstijn et al.'s study (1996) and her previous study (2010b) about the use of paper dictionaries and found a different result. She divided 60 ESL freshmen into three experimental groups: (1) reading with marginal glosses, (2) reading with paper dictionaries, (3) reading with electronic dictionaries. All three groups were given a reading comprehension passage and a surprise L2 into L1 translation task. The posttest was repeated after five days for the delayed results. Xu (2010a) found that the two dictionary groups performed better than the gloss group in both the immediate and the delayed posttests while there was little difference between the performances of the two dictionary groups.

Interestingly, the three aforementioned studies comparing marginal glosses and dictionary use found opposite results. The differences in the results can be put down on the intervening factors that can be seen in each study (i.e. the factor of need, learner proficiency,

the language if glosses and dictionaries etc.). All this illustrates the need for further research that will control the extraneous effects of the abovementioned variables.

2.5. Research Gaps and the Present Study

Although the important gaps in the literature are mentioned above as well as in the *Introduction*, it is necessary to list them here in order to walk the reader through the process that leads to the research questions before moving on to the methodology section.

Firstly, the need factor that Hulstijn and Laufer (2001a) see as indispensable was not controlled for in most of the studies discussed. Therefore, the results do not give a clear demonstration of the role of the search factor in L2 incidental vocabulary acquisition. Even though Hulstijn et al. (1996) gave the participants a reason to understand the passage well, they did not ensure that the reading task required the knowledge of the target words from the participants. Indeed, in an earlier study, Hulstijn (1993) looked at when students are more inclined to consult dictionaries during a reading task and found that the more relevant word was, the more it was looked up, which is also supported by a more recent empirical study by Peters (2007). As the need factor is closely related to motivational factors (Hulstijn & Laufer, 2001a; Xu 2010a), a study looking at the effect of the search factor when the need factor is controlled for must make sure that the reading task given to the participants requires the knowledge of the target words.

Secondly, lexical inference is a key issue when vocabulary acquisition is aimed at (Hamada, 2009; Pulido, 2007). In order to ensure that the results of the posttests are due to the use of glosses or dictionaries, the meanings of the target words should not be inferred from the passage. In this regard, the passage should be modified carefully in a way that it

gives the participants no clues about the meanings of the target words. Due to the same reason, the passage should not be about a topic that the participants are familiar with (Pulido, 2007). None of the studies controlled the lexical inference use.

Thirdly, considering that the results of the study might be affected by the language in which definitions of the words are offered (Cheng, 2009; Jacobs, Dufon & Hong, 1994; Farvardin & Biria, 2012; Ko, 2012; Xu, 2010a), both the glosses and the dictionaries should be in the same language, which was not the case in previous studies.

In addition to all this, one important limitation seen in all of the aforementioned studies on marginal glossing and dictionary use is that they all evaluate the participants' performance through tests that measure their receptive knowledge of meaning only. However many researchers agree that knowing a word requires a lot more than mere knowledge of the meaning (Aitchison, 1994; Laufer, 1998; Miller, 1999; Nation, 1990, 2001; Webb, 2005, 2007). Although the literature has many studies looking into the acquisition of different receptive and productive word aspects through various tasks, the lack of studies measuring the effects of glosses and dictionary use on learners' productive knowledge of the target words, which is also pointed out as a gap in the literature by Ko (2012), is left as one question to be answered in the realm of incidental vocabulary learning along with the exploration of different aspects of the vocabulary knowledge gained by the learners.

The present study was designed in a way to control the above-mentioned variables (i.e. lexical inference, the factor of need, language of definitions) and also to broaden the scope of the previous studies by exploring different aspects of incidental vocabulary acquisition through four different posttests: (1) receptive knowledge of meaning, (2) receptive knowledge of word class, (3) productive knowledge of meaning, and (4) productive knowledge of word class. These four features were chosen as the focus for this study.

Chapter III

2. Research Design and Methodology

3.1. Research Design

The study is a quantitative, quasi-experimental one with a pre-test and two sets of posttests. A pure experimental study could not be carried out as the students that the researcher used as the participants of her study were already assigned to classes and the researcher could not divide them into new groups. The study has no hypotheses due to the limited amount of research comparing the efficacy of L2 glosses and online monolingual dictionaries, which gives it an exploratory nature. The findings of the study will hopefully not only serve the literature but also have crucial pedagogical implications. Therefore, it is best to consider it as a study in applied linguistics.

3.2. Participants

The study was carried out in the English Language Institute of a private university in Egypt, where the researcher works as a teaching fellow. This university was chosen as the place of the study for convenience reasons. The students of this university are usually considered ESL learners due to the prevailing use of English on the campus even though this is not the case off campus.

Participants of the study were 57 ENGL100-level IEP students who had taken the TOEFL or the IELTS test in order to be accepted to the program and had all passed the cut-score for acceptance, which was 76-82 for TOEFL IBT and 6 for IELTS. Therefore, the level of the students can be described as upper-intermediate. The study was carried out during the

third, fourth and fifth weeks of the ENGL100 classes in the spring 2013 semester. Even though the students were taking reading, vocabulary, grammar, and writing classes at the time of the experiment, their levels were thought to be more or less the same as when they first entered the program as two weeks is an amount of time that is too limited to allow a considerable change in their level.

Five ENGL100 classes taught by five different ELI teachers were informed about the details of the study after the IRB permission and necessary permissions from the director of the program were taken. The students were falsely informed that the study measured their reading skills in order to avoid any unwanted focus on the vocabulary items. Every participant was given a chocolate bar for their participation. The participants were sent e-mails summarizing the real purpose of the study as well as the results and the implications of the findings after the data collection procedure.

3.3. Data Collection Procedures

The first two ENGL100 classes were used as the gloss group and the third and fourth classes were put in the dictionary group. The fifth class was divided into two and each sub-group was included in one of the two experimental groups in order to reach a similar number of students in each group. The data collection procedure followed the order given below.

3.3.1. Pretest

Both treatment groups were first given a pretest in their own classrooms. The pretest was a receptive test of recognition (see *Appendix I*). The test included a total number of 30 words. The eight target words of the study were listed along with the 2 back-up target words and 20 other English words. The students were asked to read the list and mark the words that

they knew and provide an L1 or L2 definition/synonym for each word they mark as known. A definition for each known word was required in order to avoid any false reports. At this point, the students were not asked to write their names on the papers to set them at ease and encourage them to take risks. For the same purpose, the students were informed that they would not be graded for this test and that it was done to see the vocabulary sizes of the ENGL100 students in general.

The pretest results showed that some of the participants had some previous knowledge of one of the words that were planned to be the target ones. That word was replaced by one of the back-up words that was not marked as known by any of the test takers. Therefore, only the words that were not marked as known by anyone were chosen as the target words. Some students marked some words as known but provided false definitions. These words were also excluded from the study as the students thought that they knew the word and this might have been misleading for them in the during-test.

3.3.2. During-test

One week later, the students were given the treatment. The dictionary group was taken to a CALL lab while the gloss group was given the treatment in their own classrooms as they did not need computers for the completion of the tests. Both groups were given the same reading passage, which was 540 words long (see *Appendix II*). They were asked to read the passage and answer some multiple choice comprehension questions which definitely required the knowledge of the target words (see *Appendix II*). They were informed that they were going to be graded for the test and that the test was aimed to show their reading comprehension skills in order to avoid the use of any possible individual vocabulary learning strategies by the participants. The first group had access to the same online dictionary (*Oxford Advanced Learner's American Dictionary Online*), which provides users with (1) L2

definitions, (2) L2 synonyms, (3) pronunciation help, (4) sample sentences, (5) syntactic class information, and (6) collocations (depending on the entry). The second group had L2 marginal glosses, which only provided L2 meanings for the target words.

3.3.3. Posttests

The participants were given a set of surprise vocabulary tests immediately after the treatment. As mentioned before, IVA considers vocabulary acquisition as a by-product of other tasks. For that reason, participants were not be informed of the vocabulary test in advance in order to avoid any special focus they might have on the vocabulary items (see *Appendix III*). The test consisted of four sections, which are explained below:

1. Pmeaning: Participants were given L2 definitions of the target words and were asked to provide the target words themselves. As this is not a spelling test, the participants were given full marks as long as the word could be recognized.

2. Rmeaning: Participants were given the target words and were asked to write a synonym or an L1/L2 definition for each.

3. Pclass: Participants were given the target words and were asked to write a sentence with each. The sentences were assessed on syntactic accuracy only and no attention was given to the meaning or the grammar.

4. Rclass: Participants were given a multiple choice test in which they were asked to choose the correct syntactic use of each target word.

The order of the four tests was decided in a way that none of the tests helped participants with the following ones. The same posttests were given to the participants one week later for delayed scores. The only difference was made in the order of the items on the tests.

3.4. Data Analysis and Treatment of the Research Questions

The two conditions that the two levels of the independent variable (marginal glosses and dictionary use) create were realized by using two treatment groups. The main research question, which is divided into two sub-questions, was answered with the results of the four posttests that were administered twice for immediate and delayed learning gains. Each correct answer was given one point. The results of the posttests were analyzed with MANOVA as there is more than one dependent variable in the study and running multiple ANOVAs would be unnecessary. The following chapters report on the data analysis and the results.

Chapter IV

4. Results

This study aimed to measure the effects of the use of marginal L2 glosses vs. online monolingual dictionaries during reading comprehension tasks on the incidental acquisition of four different vocabulary features. It had a pretest/posttest design. Two separate one-way MANOVAs were conducted to analyze the two groups' immediate and delayed performances on (1) receptive tests of meaning and word class and (2) and productive tests of meaning and word class. The results are reported below with the necessary tables illustrating the statistical information.

4.1. Results of the Immediate Tests

Table 1 shows the means, standard deviations, and the number of participants for the immediate tests in each group and for all four dependent variables.[1] The mean scores (the mean of the correct responses out of eight questions in each test) show a superiority of the online dictionary group over the gloss group on all four tests. This superiority is clearer in the mean scores of the productive tests than in those of the receptive ones. It is also noticed that both groups performed better in word class knowledge tests (both productive and receptive) compared to meaning tests.

The highest scores were achieved by the online dictionary group on the productive knowledge of the word class (Pclass) test ($M = 6.11$, $SD = 1.7$) and the lowest scores were

---

[1]   For individual test scores of the participants, please see *Appendix IV*.

obtained by the marginal gloss group on the productive knowledge of meaning test ($M =$ 1.14, $SD = 1.2$).

Table 1

Descriptive Statistics (Immediate test)

Vocabulary aid type		Mean	Std. Dev.	N
Pmeaning	Marginal Gloss Group	1.14	1.297	28
	Online Dictionary Group	3.07	1.730	27
	All	2.09	1.798	55
Rmeaning	Marginal Gloss Group	4.79	2.315	28
	Online Dictionary Group	5.48	2.276	27
	All	5.13	2.302	55
Pclass	Marginal Gloss Group	5.04	1.644	28
	Online Dictionary Group	6.11	1.717	27
	All	5.56	1.751	55
Rclass	Marginal Gloss Group	5.46	1.575	28
	Online Dictionary Group	5.81	1.841	27
	All	5.64	1.704	55

While the descriptive statistics show that the online dictionary group performed better than the marginal gloss group on all tests, it is necessary to look at the results of one-way MANOVA to see if this superiority is a significant one. *Table 2* shows that there is a statistically significant difference between the performances of the two groups on the immediate tests of productive knowledge. Significant differences were found on the productive test of meaning, F (4,50), $p < .001$ and on the productive test of word class, F (4,50), $p < .05$.

Table 2

Results of MANOVA for Immediate Tests

Effect		Value	F	Hypothesis df	Error df	Sig.
Intercept	Wilks' Lambda	.058	202.458[b]	4.000	50.000	.000
Group	Wilks' Lambda	.607	8.105[b]	4.000	50.000	.000

Note: a. Design: Intercept + Group, b. Exact statistic, Computed using alpha = .05

The between-subjects effect is demonstrated in *Table 3*. Based on the results of *Table 3*, it can be concluded that both groups performed statistically similarly on both receptive tests. In other words, neither group made higher learning gains than the other in receptive knowledge of meaning and word class, $F_{(4,50)}$, $p = .266$, $F_{(4,50)}$, $p = 451$, *respectively*.

Table 3

Between Subjects Effects for Immediate Tests

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	Pmeaning	244.429	1	244.429	105.083	.000
	Rmeaning	1448.981	1	1448.981	274.806	.000
	Pclass	1707.896	1	1707.896	604.945	.000
	Rclass	1748.671	1	1748.671	597.785	.000
Group	Pmeaning	51.265	1	51.265	22.040	.000
	Rmeaning	6.654	1	6.654	1.262	.266
	Pclass	15.896	1	15.896	5.631	.021
	Rclass	1.689	1	1.689	.577	.451
Error	Pmeaning	123.280	53	2.326		
	Rmeaning	279.455	53	5.273		
	Pclass	149.631	53	2.823		
	Rclass	155.038	53	2.925		
Total	Pmeaning	415.000	55			
	Rmeaning	1732.000	55			
	Pclass	1868.000	55			
	Rclass	1904.000	55			

Note: a. R Squared = .294 (Adjusted R Squared = .280), b. R Squared = .023 (Adjusted R Squared = .005), c. R Squared = .096 (Adjusted R Squared = .079), d. R Squared = .011 (Adjusted R Squared = -.008), Computed using alpha = .05

4.2. Results of the Delayed Tests

The delayed scores showed a similar superiority of the online dictionary group over the marginal gloss group on all four tests. *Table 4* shows the means, standard deviations, and the number of participants for the delayed tests in each group and for all four dependent variables. The highest scores were achieved by the online dictionary group on the productive knowledge of word class test ($M = 6.07$, $SD = 1.328$) while the lowest scores were obtained by the marginal gloss group on the productive knowledge of meaning ($M = 2.14$, $SD = 1.433$).

Table 4

Descriptive Statistics (Delayed tests)

Vocabulary Aid type		Mean	Std. Deviation	N
D-Pmeaning	Marginal Gloss Group	2.14	1.433	28
	Online Dictionary Group	2.52	1.784	27
	All	2.33	1.611	55
D-Rmeaning	Marginal Gloss Group	4.54	1.478	28
	Online Dictionary Group	5.11	1.928	27
	All	4.82	1.722	55
D-Pclass	Marginal Gloss Group	5.29	1.213	28
	Online Dictionary Group	6.07	1.328	27
	All	5.67	1.320	55
D- Rclass	Marginal Gloss Group	5.43	1.620	28
	Online Dictionary Group	5.89	1.601	27
	All	5.65	1.613	55

Table 5 shows the result of the one-way MANOVA, which revealed that this superiority of the online dictionary group over the gloss group was statistically significant only on the Pclass test, $F_{(4, 50)}$, $p < .05$. It was also seen that the difference between the performances of the two groups on Pmeaning test was no longer a significant one week later, $F_{(4, 50)}$, $p = .392$.

Table 5

Tests of Between-Subjects Effects for Delayed Tests

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	D-Pmeaning	298.667	1	298.667	114.565	.000
	D-Rmeaning	1279.169	1	1279.169	435.620	.000
	D-Pclass	1773.779	1	1773.779	1098.686	.000
	D-Rclass	1760.585	1	1760.585	678.508	.000
Group	D-Pmeaning	1.940	1	1.940	.744	.392
	D-Rmeaning	4.551	1	4.551	1.550	.219
	D-Pclass	8.543	1	8.543	5.292	.025
	D- Rclass	2.913	1	2.913	1.122	.294
Error	D-Pmeaning	138.169	53	2.607		
	D-Rmeaning	155.631	53	2.936		
	D-Pclass	85.566	53	1.614		
	D-Rclass	137.524	53	2.595		
Total	D-Pmeaning	438.000	55			
	D-Rmeaning	1437.000	55			
	D-Pclass	1864.000	55			
	D- Rclass	1899.000	55			

Note: a. R Squared = .014 (Adjusted R Squared = -.005), b. R Squared = .028 (Adjusted R Squared = .010), c. R Squared = .091 (Adjusted R Squared = .074), d. R Squared = .021 (Adjusted R Squared = .002)

4.3. Loss of Vocabulary Knowledge

Figure 1 shows the differences between the mean scores of each immediate and delayed test taken by both groups. The performances of both groups deteriorated in all tests after one week except for the performance of the gloss group on the productive knowledge of the meaning and word class tests.

Figure 1

Loss of Vocabulary Knowledge after One Week
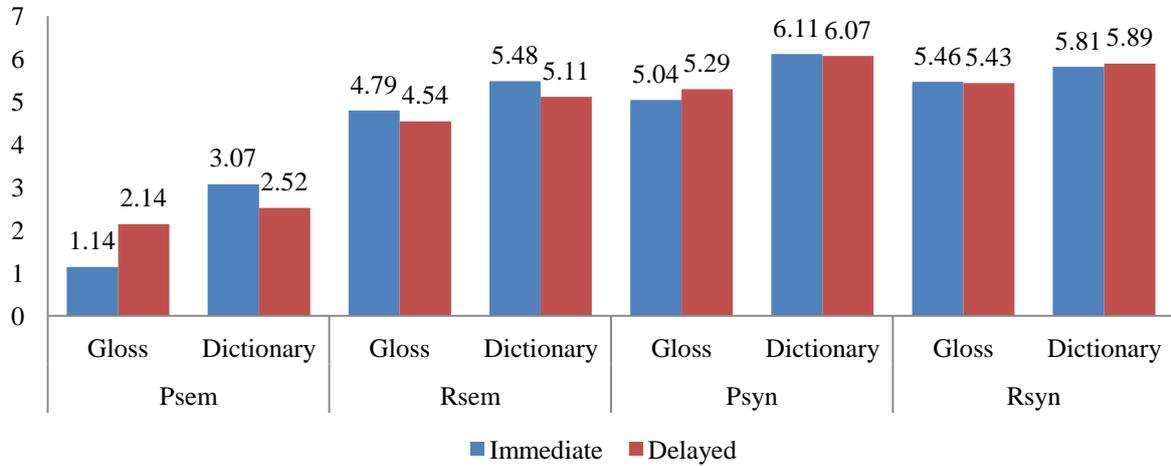
Both groups lost between 0.5% and 17% of their vocabulary learning gains one week after the immediate tests. The online dictionary group lost their gains of productive knowledge of meaning by 17%, receptive knowledge of meaning by 6.7%, productive knowledge of word class by 0.6% and receptive knowledge of word class by 1.5%. . The gloss group lost their gains of receptive knowledge of meaning by 5.2%, and receptive knowledge of word class by 0.5%. After having taken the immediate tests, the gloss group improved their productive knowledge of meaning by 87%, productive knowledge of word class by 4.9%, the possible reasons for which is discussed in the following chapter.

Chapter V

5. Discussion and Conclusion

The results of the study show a statistically significant difference between the performances of the gloss group and the dictionary group on the tests of productive knowledge of meaning and word class while no difference is observed between their performances on the receptive knowledge of meaning and word class tests. The delayed scores indicate that the productive knowledge of word class test is the only one that shows a considerable difference between the performances of the two groups one week after the treatment.

5.1. Receptive Knowledge of Meaning and Word Class

In their 1996 study, Hulstijn et al. saw that the paper dictionary group outperformed the gloss group when dictionaries were used properly. Xu (2010a), on the other hand, found that the gloss group performed better than the paper dictionary group. Interestingly, in her later study (2010b), she noted that learners who used electronic dictionaries had higher learning gains than their peers who were provided with glosses.

As opposed to the findings of the above-listed studies, no statistically significant difference was observed between the performances of the two groups of this study either in the immediate or the delayed receptive tests. The results of the receptive meaning tests contradict with the findings of Xu (2010a; 2010b) and Hulstijn, et al. (1996), who both noted that one of the two groups outperformed the other on receptive knowledge of meaning tests. This difference between the findings of the present study and those of previous studies can be attributed to the fact that none of the previous studies controlled for the variables that are strictly controlled for in the present study (e.g. lexical inference, topic familiarity,

morphological cues, the need factor). Also, none of the previous studies used online dictionaries, whose effect might not be the same as that of other dictionary types.

The results of the receptive knowledge of word class test also show that the two treatment groups performed similarly. The knowledge of word class was not tested in any of the previous studies looking at the comparative efficacy of glosses and dictionaries. Therefore it is not possible to compare the results to any other findings in the literature. However, that the mean scores for word class tests are always higher than those for meaning tests suggests that meaning is not the earliest acquired feature of a word (*see Table 1*). Indeed, Pellicer-Sanchez and Schmitt (2010) noted that the incidental acquisition of word class took place earlier than that of meaning in their study, too. The present study showed that the order of incidental acquisition of different word features after one encounter during a reading comprehension task was Rclass, Pclass, Rmeaning, and Pmeaning, word class knowledge being acquired earlier than meaning.

The delayed scores indicated that students' vocabulary retention (receptive) also benefited from online dictionaries and marginal glosses to the same extent. Both groups had little or no loss of their knowledge of the target items one week after the immediate tests. However, it is noteworthy to mention that the results of the delayed tests are not only due to the treatment but also to the practice the participants have through the immediate tests. Therefore, the scores would probably be lower if no immediate tests had been involved in the process.

Overall, the results of the receptive tests do not support Hulstijn and Laufer's (2001) involvement load theory as no significant difference was seen in the receptive performances of the two groups either in immediate or delayed tests.

5.2. Productive Knowledge of Meaning and Word Class

Unlike the receptive tests, the productive tests revealed a significant difference between the performances of the two groups. The dictionary group outperformed the gloss group on both immediate productive tests (Pmeaning and Pclass). As none of the three studies on the comparative effects of dictionaries and glosses on IVA looked at learners' productive knowledge of the target words, it is not possible to compare the results of the two productive posttests to other studies in the literature. However, the findings of the present study for the productive tests reflect that dictionary use requires deeper processing of the target words compared to marginal glosses as productive knowledge is reported to be achieved after receptive knowledge by many researchers. While the results of the receptive tests did not show any difference between the efficacies of the two treatments, productive tests demonstrated that marginal glosses failed to encourage productive knowledge as much as the dictionary use did.

Hulstijn and Laufer's (2001) involvement load theory is supported by the findings of the study on the productive level as dictionary use, which, they argue, requires more involvement load as it includes both the need and the search factors, resulted in better acquisition of vocabulary than marginal gloss use. Therefore, more involvement load resulted in better acquisition.

The delayed scores reveal that this significant difference between the performances of the two groups remains significant only on the Pclass test. Even though the delayed scores did not show any significant differences between the groups in terms of knowledge of meaning, there was a clear difference between the performances of the two groups in terms of the knowledge of word class.

The changes in the scores between the immediate and delayed tests are surprising as the gloss group improved their performance on the Pmeaning and Pclass. The mean scores for the delayed Pmeaning and Pclass are higher than the immediate Pmeaning and Pclass scores, which might seem contradictory at first considering that immediate scores are usually expected to be higher than delayed ones. However, studies using immediate and delayed vocabulary tests on the same target items tend to have such an increase in scores as the immediate tests serve as additional practice for the participants and some perform better in the delayed test (Webb, 2007). Although the marginal gloss group seems to have improved their productive knowledge of meaning and word class, their performances on these two tests are not as high as those of the dictionary group, which shows that the improvement is not a significant one and that it can be attributed to the fact that they practice their productive skills in the immediate tests and therefore perform slightly better in the delayed productive tests.

It is clear that dictionary use results in deeper processing of vocabulary by learners compared to gloss use although the two techniques do not differ in their efficacy when basic processing skills such as receptive knowledge of meaning are aimed at. This suggests that if glosses and dictionaries are compared for their efficacy in facilitating the acquisition of other features of word knowledge (i.e. association of meaning and orthography), significant differences in the results can be discovered in future research.

5.3. Pedagogical Implications

Learners can indeed acquire vocabulary intentionally after one encounter as the study indicates. The results show that both marginal glosses and online dictionaries are useful vocabulary aiding tools when receptive vocabulary knowledge is aimed at. However, when productive vocabulary knowledge is the goal, online dictionaries provide learners with higher learning gains both on meaning and word class knowledge of vocabulary. Therefore,

dictionary use should be encouraged inside and outside classrooms as it has clear advantages to learners compared to marginal glosses.

However, incidental vocabulary acquisition after one encounter and with the help of either online dictionaries or marginal glosses has its limits as well. The delayed scores show that both groups tend to lose their learning gains to a certain extent. This shows that neither glosses nor online dictionaries are enough for retention and further practice and more frequent encounters are needed, which is a reminiscent of Nation's emphasis (2001) on the necessity of multiple encounters with a word for its successful acquisition and the findings of Webb's 2007 study, showing that learners have considerable incidental learning gains of a word only after seeing it at least 10 times. Therefore, it is possible to see online dictionaries and marginal glosses as tools that constitute the first step towards learning a word, the former being a more effective one than the latter.

Lastly, it should be noted that the present study did not create an ideal learning environment for the participants as it controlled for many variables (e.g. topic familiarity, lexical inference, etc.) that learners generally use in their authentic learning situations. This was done in order to ensure that the results showed how effective marginal glosses and online dictionaries are as learning tools per se without any extraneous effects of these variables that are usually present in the process of IVA. Teachers should take into consideration how effectively these two vocabulary aiding tools are utilized by the students.

5.4. Suggestions for Future Research

The present study looked at how online dictionaries and marginal glosses affect receptive and productive knowledge of meaning and word class. The results showed that when the posttests are modified in away to measure individual aspects of vocabulary,

significant differences are seen. Future research can investigate the effects of the two vocabulary aids on the acquisition of other vocabulary aspects such as orthography, association, and pronunciation. Also, it would be very interesting to see how participants perform in other types of posttests such as fill-in-the-blanks, meaningful sentence writing, etc. Future research can also look at how individual participants perform on each separate test after the treatment as well as to look at how learners of different levels benefit from the treatment. Lastly, future studies can use a design that requires the participants to take an immediate test on some words and a delayed test on others, which would eliminate the positive effect of having the immediate tests on the scores of the delayed one.

5.5. Conclusion

Print dictionaries were once thought to be a staple of every language learner's library. It is easy to see now that they are replaced by more practical vocabulary aiding tools such as online dictionaries, electronic dictionaries, hypertext glosses, marginal glosses and many more. Interestingly, very few studies have compared the efficacy of these techniques on vocabulary acquisition despite the fact that having an effective encounter with a previously unknown word can create a stronger base for additional practice on the word as well as increase the likelihood of remembering the word on later encounters. This study looked at the comparative efficacy of two of these common vocabulary tools and found important results. It is hoped that future studies will extent the scope of this study and compare these tools on other aspects of vocabulary acquisition and enable learners to resort to the most effective of these vocabulary aiding tools for more effective first vocabulary encounters.

References

Ahmed, J. (2011). Intentional vs. incidental vocabulary learning. *Interdisciplinary Journal of Contemporary Research in Business*, *3* (5), 67-77.

Aitchison, J. (1994). *Words in the Mind*. Oxford: Blackwell.

Alemi, M., & Tayebi, A. (2011). The influence of incidental and intentional vocabulary acquisition and vocabulary strategy use on learning L2 vocabularies. *Journal of Language Teaching and Research*, *2* (1), 81-98.

Allesi, S., & Dwyer, A. (2008). Vocabulary assistance before and during reading. *Reading in a Foreign Language* , 20 (2) 246–263.

Aust, R., Kelley, M. J., & Roby, W. B. (1993). The use of hyper-reference and conventional dictionaries. *Educational Technology Research and Development, 41*, 63–73.

Barcroft, J. (2004). Second language vocabulary acquisition: a lexical input approach. *Foreign Language Annals, 37*, 200–208. doi:10.1111/j.1944-9720.2004.tb02193.x.

Bruton, A. , Lopez, M.G. , & Mesa, R. E. (2011). Incidental L2 vocabulary learning: an impractible term? . *TESOL Quarterly*, *45* (4), 758-768.

Chen, S. (2008). Monolingual, bilingual dictionaries and language study. *Cross-cultural Communication ,4* (1), 65-73.

Chen, Y. (2010). Dictionary use and EFL learning: a contrastive study of pocket electronic dictionaries and paper dictionaries'. *International Journal of Lexicography, 23* (3), 275-306.

Chen, Y. (2011). Dictionary use and vocabulary learning in the context of reading. *International Journal of Lexicography*, *24, na*.

Cheng, Y.H. (2009). L1 glosses: effects on EFL learners' reading comprehension and vocabulary retention. *Reading in A Foreign Language, 21* (2), 119-142.

Cho, K. S., & Krashen, S. D. (1994). Acquisition of vocabulary from the sweet valley kids series: adult ESL acquisition. *Journal of Reading, 37* (8), 662-667.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: a framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*, 671–684.

Crow, J. (1986). Receptive vocabulary acquisition for reading comprehension. *The Modern Language Journal, 70* (3), 242-250.

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

Dziemianko A. (2010). Paper or electronic?: the role of dictionary form in language reception, production and the retention of meaning and collocations. *International Journal of Lexicography, 23* (3), 257-274.

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

Ellis, R., & He, X. (1999). The roles of modified input and output in the incidental acquisition of word meanings. *Studies in Second Language Acquisition, 21*, 285–301.

Fan, M. (2000). How big is the gap and how to narrow it? An investigation into the active and passive vocabulary knowledge of L2 learners. *RELC Journal*, *31*, 105–119

Farvardin, M. T., & Biria, R. (2012). The impact of gloss types on Iranian EFL students' reading comprehension and lexical retention. *International Journal of Instruction*, *5* (1), 99-116.

Folse, K. S. (2006). The effect of type of written exercise on L2 vocabulary retention. *TESOL Quarterly, 40*, 273–293.

Fraser, C.A. (1998). The role of consulting a dictionary in reading and vocabulary learning. *The Canadian Journal of Applied Linguistics, 2*, 73-90.

Gass, S. (1999). Discussion: incidental vocabulary acquisition. *Studies in Second Language Acquisition, 21 (2),* 319-333.

Gouws, R.H. (2004). Monolingual and bilingual learners' dictionaries. *Lexikos, 14*, 264-274

Hamada, M. (2009). Development of L2 word-meaning inference while reading. *System: An International Journal of Educational Technology and Applied Linguistics, 37*, 447-460.

Hu, M., & Nation, I. S. P. (2000). Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language, 13*, 403–430.

Hulstijn, J. (1990). A comparison between the information-processing and the analysis / control approaches to language learning. *Applied Linguistics*, *11*, 30–45.

Hulstijn, J. H. (1992). Retention of inferred and given word meanings: Experiments in incidental word learning. In P. J. L. Arnaud & H. Béjoint (Eds.), *Vocabulary and Applied Linguistics* (113–125). London: Macmillan.

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

Hulstijn, J.H. , Hollander, M., & Greidanus, T. (1996). Incidental vocabulary learning by

advanced foreign language students: the influence of marginal glosses, dictionary use,

and reoccurrence of unknown words. *The Modern Language Journal*, *80* (3), 327-339.

Hulstijn, J. H., & Laufer, B. (2001a). Some empirical evidence for the involvement load

hypothesis in vocabulary acquisition. *Language Learning*, *51*, 539–558.

Hulstijn, J. & Laufer, B. (2001b). Incidental vocabulary acquisition in a second language: The

construct of task-induced involvement. *Applied Linguistics*, *22* (1), 1-26.

Holly, F., & King, J. (1971). Vocabulary glosses in foreign language reading materials.

Language Learning, *21*, 213–219. doi:10.1111/j.1467-1770.1971.tb00060.x

Huckin, T & Coady, J. (1999). Incidental vocabulary acquisition in a second language: A review.

Studies of Second Language Acquisition 21, 181-193.

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.

Jian, H., Sandnes, F. E., Law, K. M. Y., Huang, Y. P., & Huang, Y. (2009). The role of

electronic pocket dictionaries as an English learning tool among Chinese students.

Journal of Computer Assisted Learning, *25*, 503–514

Kaivanpanah, S., & Alavi, M. (2008). Deriving Unknown Word Meaning from Context : Is it

Reliable?. *RELC Journal* , *39*(1) 77-95.

Keating, G. D. (2008). Task effectiveness and word learning in a second language: The

involvement load hypothesis on trial. *Language Teaching Research, 12* (3), 365–86.

Krashen, S. (1985). *The input hypothesis: Issues and implications*. Harlow, UK: Longman.

Krashen, S. (1989). We acquire vocabulary and spelling by reading: additional evidence for the

    input hypothesis. *The Modern Language Journal, 73*, 440-64.

Knight, S. (1994). Dictionary: the tool of last resort in foreign language reading?: a new

    perspective. *Modern Language Journal, 78* (3), 285-299.

Ko, M. H. (2005). Gloss, 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.

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

    you think you know, and words you can't guess. In J. Coady & T. Huckin (Eds.), *Second*

    Language Vocabulary Acquisition (pp. 20–34). New York, NY: Cambridge University

    Press.

Laufer, B. (1998). The development of passive and active vocabulary: same or different? *Applied*

    Linguistics, *19*, 255–271.

Laufer, B. (2011). The contribution of dictionary use to the production and retention of

    collocations in a second language. *International Journal of Lexicography, 24* (1), 29–49.

Laufer, B., & Hadar, L. (1997). Assessing the effectiveness of monolingual, bilingual, and

    "bilingualised" dictionaries in the comprehension and production of new words. *The*

    Modern Language Journal, 81(2), 189-196.

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 & Levitzky-Aviad (2006). Examining the effectiveness of 'bilingual dictionary plus: a

    dictionary for production in a foreign language. *International Journal of Lexicography*,

    19 (2), 135-156.

Laufer, B. & Hill, M. (2003). Type of task, time-on-task and electronic dictionaries in incidental

    vocabulary acquisition. *International Review of Applied Linguistics in Language*

    Teaching, 41 (2), 87-106.

Laufer, B. and Yano, Y. (2001). Understanding unfamiliar words in a text: Do L2 learners

    understand how much they don't understand? *Reading in a Foreign Language, 13*(2),

    549-66.

Lew, R. (2011). Studies in dictionary use: recent developments. *International Journal of*

    Lexicography, 24 (1), 1-4.

Liu, J. (2008). The effect of lengthened writing approach on L2 vocabulary acquisition. *Asian*

    Social Science, *4* (4), 129-134.

Liu, N. & Nation, I.S.P. (1985). Factors affecting guessing vocabulary in context. *RELC Journal,*

    16, 33-42.

Lu, M. (2008), Effectiveness of vocabulary learning via mobile phone. *Journal of Computer*

    Assisted Learning, 24, 515-525.

Luppescu, S., & Day, R. R. (1993). Reading, dictionaries, and vocabulary learning. *Language*

    Learning, 43 (2), 263-287.

Meara, P. (1990). A note on passive vocabulary. *Second Language Research, 6* (2), 150-154.

Miller, G. A. (1999). On knowing a word. *Annual Review of Psychology*, *50*, 1–19.

Mondria, J. (2003). The effects of inferring, verifying, and memorizing on the retention of L2 word meanings. *Studies in Second Language Acquisition*, *25*, 473-499.

Mondria, Y., & Wiersma, B. (2004). 'Receptive, productive and receptive + productive L2 vocabulary learning: what difference does it make?' In P. Bogaards & B. Laufer (Eds.), *Vocabulary in a Second Language (*155-69 *)*. Amsterda: John Benjamins.

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

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

Nation, P. (2001) *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press.

Nation, P. (2006). How large a vocabulary is needed for reading and listening? *The Canadian Modern Language Review, 63*, 59–82.

Oxford, D. , & Crookall, R. (1990). Learning vocabulary: a critical anaylsis of techniques. *TESL Canada Journal*, *7* (2), 9-31.

Paribakht, T.S. (2005). The influence of First language lexicalization on second language lexical inferencing: A study of Farsi-speaking learners of English as a foreign language. *Language Learning*, *55*, 701-748

Paribakht, T. S., & Wesche, M. (1999). Reading and "incidental" L2 vocabulary acquisition. *Studies in Second Language Acquisition*, *21*, 195–224.

Pellicer-Sánchez, A., & Schmitt, P. (2010). Incidental vocabulary acquisition from an authentic novel: Do Things Fall Apart?. *Reading in a Foreign Language, 22* (1), 31-55.

Peters, E. (2007). Manipulating L2 learners' online dictionary use and its effect on L2 word retention. *Language Learning and Technology, 11*, 36–58.

Pichette, F., De Serres, L. & LaFontaine, M. (2012), Sentence reading and writing for second language vocabulary acquisition. *Applied Linguistics*, *33* (1), 66-82.

Pigada, M., & Schmitt, N. (2006). Vocabulary acquisition from extensive reading: A case study. *Reading in a Foreign Language, 18*, 1–28.

Pulido, D. (2004). The effect of cultural familiarity on incidental vocabulary acquisition through reading . *The Reading Matrix: An Online International Journal*, *4,* 20-53.

Pulido, D. (2007). The effects of topic familiarity and passage sight vocabulary on L2 lexical inferencing and retention through reading. *Applied Linguistics*, *28,* 66-86.

Qian, D. (2004) Second language lexical inferencing: preferences, perceptions and practices. In P. Bogaards and B. Laufer (eds.), *Vocabulary in a Second Language*. (155-69) Amsterdam: John Benjamins.

Richards, J. C. (1976). The role of vocabulary teaching. *TESOL Quarterly*, *10* (1), 77-89.

Read J. (2000) *Assessing Vocabulary*. Cambridge University Press, Cambridge, UK.

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

Schmitt, N. (1997). Vocabulary learning strategies. In N. Schmitt & M. McCarthy (Eds.) *Vocabulary: description, acquisition and pedagogy* (199-227). Cambridge: Cambridge University Press.

Schmitt, N., & Schmitt, D. (1995). Vocabulary notebooks: theoretical underpinnings and practical suggestions. *ELT Journal, 49*, 113-142.

Shahrokni, S. A. (2009). Second language incidental vocabulary learning: The effect of online textual, pictorial, and textual pictorial glosses. *TESL-EJ*, *13*(3).

Stoddard, G. D. (1929). An experiment in verbal learning. *Journal of Educational Psychology*, *20*(6, 452-457. doi:10.1037/h0073293.

Swain, M. (1985). Communicative competence: some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds): *Input in Second Language Acquisition.* (235–53). Newbury House.

Watanabe, Y. (1997). Effects of single and multiple-choice gloss on incidental vocabulary learning. *JACET Bulletin*, *28*, 177-191.

Webb, S. (2005). Receptive and productive vocabulary learning: the effects of reading and writing on word knowledge. *Studies in Second Language Acquisition, 27*, 33–52.

Webb, S. (2007). The effects of repetition on vocabulary knowledge. *Applied Linguistics, 28* (1), 46–65.

Webb, S. (2008). Receptive and productive vocabulary sizes of l2 learners. *SSLA*, 30, 79-95.

White, P. A. (1997). The Role of Dictionaries in Language Learning, *Dialog on Language Instruction, 12,* 7-16.

Wode, H. (1999). Incidental vocabulary acquisition in the foreign language classroom. *Studies in Second Language Acquisition*, *21*, 243–258. doi:10.1017/S0272263199002053

Xu, J. (2009). An experimental study on the effects of different reading tasks on L2 vocabulary acquisition. *English Language Teaching, 2* (3), 69-79.

Xu, X. (2010a). An empirical study on the effect of task on L2 incidental vocabulary acquisition through reading. *Asian Social Science*, *6*(7).

Xu, X. (2010b). Study on the effect of dictionary use on second language incidental vocabulary

acquisition- an empirical study of college English vocabulary learning strategy. *Journal

of Language Teaching and Research*, *1*(4), 519-523.

Yoshii, M. (2006). L1 and L2 gloss: their effects on incidental vocabulary learning. *Language

Learning & Technology, 10*(3), 85-101.

Zandieh, Z. (2012). The effects of hypertext gloss on comprehension and vocabulary retention

under incidental and intentional learning conditions. *English Language Teaching, 5* (6),

60-73.

Zhou, S. (2010). Comparing receptive and productive academic vocabulary knowledge of

Chinese EFL learners. *Asian Social Science, 6* (10), 14-19.

Appendices

Appendix I

Vocabulary Knowledge Test

*This test is designed to show IEP students' vocabulary size. You **will not** be graded for this test. Read the following words and put a mark beside the ones that you know. Please provide the meanings or synonyms of the words that you mark either in English or in Arabic.*

Senior…………………	Piety………………..	Skiff …………………
Fervid ……………	Vanish……………….	Correspond…………….
Vast……………………..	Noble ………………	Eruption………………
Coeval…………………	Limpid……………	Violent………………….
Trim………………….	Sink…………………	Massive…………………
Zone…………………	Apex…………………	Diminish………………
Lambaste ……………	Submerge……………	Eschew………………….
Vessel………………	Catastrophe……………	Infinite………………..
Anon…………………	Trace………………	Construct ……………
Bury …………………	Circa …………………	Hubris…..

Appendix II

READING TEST *(Gloss Group)*

*This reading-comprehension test is designed to measure your reading skills. There is no time limit for this test. You will be graded for this test. Each correct answer is given **1** **point**. If you need any help with the vocabulary, you can refer to the marginal glosses. Good luck!*

Atlantis

Thirty-five years ago, a tremendous explosion blew apart an island and completely destroyed a civilization called Atlantis. Where was Atlantis? What kind of people lived there? Why and how was it destroyed? No one knows the answers to these questions, but there have been hundreds of guesses and theories.

The Greek philosopher Plato is the main source for the legend of Atlantis. His description of the "lost continent" still excites the modern mind. Plato's Atlantis was a kind of paradise – a big island with wonderful mountain ranges, green plains that were full of every variety of animal, luxuriant gardens where the fruit was "fair and wondrous in infinite abundance.". The people were especially <u>fervid</u> about precious stones like diamond and amber.

passionate

The capital of Atlantis was beautifully constructed in white, black, and red stone. Plato says that the city was full of life, activity, and culture. The people travelled on beautiful <u>skiff</u>s and were very sociable.

small boats

However, an unexpected end was waiting for Atlantis. Its culture began to decay. Then, one day and one night, the continent was completely destroyed. Plato concluded that the most important factor that resulted in their downfall was their <u>hubris</u>. Plato recounts that the people changed their law-respecting way of life. They gradually began to live with less piety and to value luxuries, wealth and idleness. At the end of his account of Atlanteans, Plato <u>lambasted</u> their behavior.

excessive pride

harshly criticize

As punishment, Plato says, the gods sent "terrible fires and earthquakes" that caused Atlantis to sink into the sea <u>anon</u>.

soon

This story intrigues people so much that many have been searching for the explanation of "the lost continent" for 23 centuries. Two questions remain unanswered. Where was Atlantis, and where did it go? There are three probable locations for Atlantis: the

Azores, in the Atlantic Ocean; the Bimini Islands, in the Caribbean Sea; and the Santorini, in the Aegean Sea. However scientists and historians <u>eschew</u> giving an answer to the question about where Atlantis really was.

avoid

Plato said Atlantis existed <u>circa</u> 9,360 BC, and that its story had been passed down by poets, priests, and others. But Plato's writings about Atlantis are the only known records of its existence.

about

Romm believes Plato created the story of Atlantis to convey some of his philosophical theories. "He was dealing with a number of issues, themes that run throughout his work," he says. "His ideas about divine versus human nature, ideal societies, the gradual corruption of human society—these ideas are all found in many of his works. Atlantis was a different tool to get at some of his favorite themes." He argues that Atlantis could have never existed because Plato described the island as carefully planned – in five zones built in perfect concentric circles surrounded by an ocean which was always <u>limpid</u> no matter what season it was.

transparent

Today, many people believe the tale to be complete fiction, the creation of a philosopher's imagination used to illustrate an argument. Others believe that the story was inspired by catastrophes in ancient civilizations. Still others maintain that the story is an accurate representation of a long lost and almost completely forgotten. The disappearance of Atlantis remains a lasting mystery.

Comprehension Questions

1. Which of the following is true about Atlanteans according to the passage?
    a. They worked hard to mine precious stones.
    b. They were very knowledgeable about precious stones.
    c. They were very passionate about precious stones.
    d. They were very talented in shaping precious stones.

2. How did Antlanteans travel from one part of the island to the other?
    a. They used small boats to travel.
    b. They traveled on foot as the island was small.
    c. They had low-tech air planes.
    d. They traveled on horseback.

3. According to Plato, what was <u>the most important factor</u> that resulted in the downfall of the Atlantis civilization?
    a. Atlanteans did not believe in gods.
    b. Atlanteans spent too much money.
    c. Atlanteans had too much pride.
    d. Atlanteans were lazy people.

4. What did Plato feel about the collapse of Atlantis?
   a. He was sorry that Atlanteans changed their behaviour.
   b. He agreed that they were meant to disappear.
   c. He criticized them harshly about their behavior.
   d. He tolerated their behaviour.

5. How long did it take Atlantis to disappear after the gods sent fires and earthquakes?
   a. It disappeared after suffering for a long time.
   b. It disappeared after 2 weeks.
   c. It disappeared on the same day.
   d. It disappeared soon.

6. What do historians and scientists say about where Atlantis was?
   a. They avoid giving an answer.
   b. They do not know the answer.
   c. They differ in their answers.
   d. They fear giving an answer.

7. When did Atlantis exist according to Plato?
   a. Before 9.360 BC.
   b. After 9.360 BC.
   c. Only in the year of 9.360 BC.
   d. About 9.360 BC.

8. What does Plato say about the ocean surrounding the island?
   a. That it was transparent at all times.
   b. That it was very polluted.
   c. That it had a green-blue color.
   d. That it was always wavy.

Were you familiar with Atlantis before reading the text?   Yes / No

READING TEST *(Dictionary Group)*

*This reading-comprehension test is designed to measure your reading skills. There is no time limit for this test. You will be graded for this test. Each correct answer is given **1 point**. If you need any help with the vocabulary, you can use the online dictionaries. Good luck!*

Atlantis

Thirty-five years ago, a tremendous explosion blew apart an island and completely destroyed a civilization called Atlantis. Where was Atlantis? What kind of people lived there? Why and how was it destroyed? No one knows the answers to these questions, but there have been hundreds of guesses and theories.

The Greek philosopher Plato is the main source for the legend of Atlantis. His description of the "lost continent" still excites the modern mind. Plato's Atlantis was a kind of paradise – a big island with wonderful mountain ranges, green plains that were full of every variety of animal, luxuriant gardens where the fruit was "fair and wondrous in infinite abundance.". The people were especially fervid about precious stones like diamond and amber.

The capital of Atlantis was beautifully constructed in white, black, and red stone. Plato says that the city was full of life, activity, and culture. The people travelled on beautiful skiffs and were very sociable.

However, an unexpected end was waiting for Atlantis. Its culture began to decay. Then, one day and one night, the continent was completely destroyed. Plato concluded that it was their hubris than anything else that resulted in their downfall. Plato recounts that the people changed their law-respecting way of life. They gradually began to live with less piety and to value luxuries, wealth and idleness. At the end of his account of the Atlantis people, Plato lambasted their behavior. As punishment, he says, the gods sent "one terrible night of fire and earthquakes" that caused Atlantis to sink into the sea anon.

This story intrigues people so much that many have been searching for the explanation of "the lost continent" for 23 centuries. Two questions remain unanswered. Where was Atlantis, and where did it go? There are three probable locations for Atlantis: the Azores, in the Atlantic Ocean; the Bimini Islands, in the Caribbean Sea; and the Santorini, in the Aegean Sea. However scientists and historians eschew giving an answer to the question about where Atlantis really was.

Plato said Atlantis existed circa 9,360 BC, and that its story had been passed down by poets, priests, and others. But Plato's writings about Atlantis are the only known records of its existence.

Romm believes Plato created the story of Atlantis to convey some of his philosophical theories. "He was dealing with a number of issues, themes that run throughout his work," he

says. "His ideas about divine versus human nature, ideal societies, the gradual corruption of human society—these ideas are all found in many of his works. Atlantis was a different tool to get at some of his favorite themes." He argues that Atlantis could have never existed because Plato described the island as carefully planned – in five zones built in perfect concentric circles surrounded by an ocean which was always limpid no matter what season it was.

Today, many people believe the tale to be complete fiction, the creation of a philosopher's imagination used to illustrate an argument. Others believe that the story was inspired by catastrophes in ancient civilizations. Still others maintain that the story is an accurate representation of a long lost and almost completely forgotten. The disappearance of Atlantis remains a lasting mystery.

Comprehension Questions

1. Which of the following is true about Atlanteans according to the passage?
   a. They worked hard to mine precious stones.
   b. They were very knowledgeable about precious stones.
   c. They were very passionate about precious stones.
   d. They were very talented in shaping precious stones.

2. How did Antlanteans travel from one part of the island to the other?
   a. They used small boats to travel.
   b. They traveled on foot as the island was small.
   c. They had low-tech air planes.
   d. They traveled on horseback.

3. According to Plato, what was the most important factor that resulted in the downfall of the Atlantis civilization?
   a. Atlanteans did not believe in gods.
   b. Atlanteans spent too much money.
   c. Atlanteans had too much pride.
   d. Atlanteans were lazy people.

4. What did Plato feel about the collapse of Atlantis?
   a. He was sorry that Atlanteans changed their behaviour.
   b. He agreed that they were meant to disappear.
   c. He criticized them harshly about their behavior.
   d. He tolerated their behaviour.

5. How long did it take Atlantis to disappear after the gods sent fires and earthquakes?
    a. It disappeared after suffering for a long time.
    b. It disappeared after two weeks.
    c. It disappeared on the same day.
    d. It disappeared soon.

6. What do historians and scientists say about where Atlantis was?
    a. They avoid giving an answer.
    b. They do not know the answer.
    c. They differ in their answers.
    d. They fear giving an answer.

7. When did Atlantis exist according to Plato?
    a. Before 9.360 BC.
    b. After 9.360 BC.
    c. Only in the year of 9.360 BC.
    d. About 9.360 BC.

8. What does Plato say about the ocean surrounding the island?
    a. That it was transparent at all times.
    b. That it was very polluted.
    c. That it had a green-blue color.
    d. That it was always wavy.

Were you familiar with Atlantis before reading the text?   Yes / No

Appendix III

VOCABULARY QUIZ I

Read the definitions given below and provide <u>the English word</u> for them. There is no time limit for this test. Each correct answer is worth 1 point (8 points overall).

Passionate……………………………

Small boat …………………………………

Excessive pride ………………………………………….

Harshly criticize ………………………………….

Soon ………………………………….

Avoid……………………………………..

Around/about (the time)………………………………………….

Transparent…………………………………….

VOCABULARY QUIZ II

Read the English words below and provide the meaning for them (<u>English and Arabic definitions are accepted.</u>) There is no time limit. Each correct answer is 1 point (8 points overall).

Fervid …………………………….

Skiff …………………………………

Hubris ………………………………….

Lambaste ………………………………………

Anon ………………………………….

Eschew ………………………………………..

Circa ………………………………………….

Limpid ………………………………………….

VOCABULARY QUIZ III

Write an English sentence with each of the words given below. There is no time limit for this test. Each correct answer is 1 point (8 points overall)

Fervid

………………………………………………………………………………………………

Skiff

………………………………………………………………………………………………

Hubris

………………………………………………………………………………………………

Lambaste

………………………………………………………………………………………………

Anon

………………………………………………………………………………………………

Eschew

………………………………………………………………………………………………

Circa

………………………………………………………………………………………………

Limpid

………………………………………………………………………………………………

VOCABULARY QUIZ IV

Choose the <u>correct use</u> of each word below. Each correct answer is 1 point (8 points overall).

1. **Hubris**

   a. He was hubrissing about himself all the time.
   b. He saw himself as a perfect human being hubris any flaws.
   c. Her hubris friend kept bragging about his new car.
   d. He was disliked by many people due to his hubris.

2. **Skiff**

   a. He could hardly stand on the skiff.
   b. He skiffed along the canal.
   c. The captain moved the ship skiff the rocks,
   d. The skiff ship moved southwards.

3. **Fervid**

   a. The writer was the one who fervided the style.
   b. The writer was very fervid about this style.
   c. The writer worked on this style fervid.
   d. He was very hard-working for a fervid.

4. **Limpid**

   a. The ocean limpids under the sun.
   b. He enjoyed his limpid on the beach.
   c. The limpid sea was very relaxing.
   d. The river was flowing limpid on the rocks.

5. **Eschew**

   a. She tried to eschew talking to him in the meeting.
   b. The biggest eschew of her life was to smoke.
   c. Her eschew actions were not welcomed.
   d. She talked eschew the important points.

6.   Lambaste

   a. She had a lambaste talk to deliver.
   b. She lambasted on the new coordinator.
   c. She is famous for her long lambastes.
   d. She talked lambaste the new coordinator.

7.   Circa

   a. Historians circa the birth of the writer in the 19th century.
   b. The writer is thought to be born circa the 19th century.
   c. The circa date of the writer's birth is thought to be the 19th century.
   d. The writer's birth date is just a circa and has no evidence.

8.   Anon
   a. An anon meeting is waiting for me.
   b. The anon was not known.
   c. I will be at the meeting anon.
   d. She anons the meeting.

Appendix IV

Table A

Individual Test Scores of the Participants

Partici pant	Group (Gloss:1 / Dict: 2)	Imm. Pmeani ng	Immed. Rmeani ng	Immed. Pclass	Immed. Rclass	Del. Pmean ing	Del. Rmean ing	Del. Pclass	Del. Rclass
1	1	2	8	6	5	5	7	7	8
2	1	4	8	7	6	5	8	6	6
3	1	1	5	5	7	1	4	3	5
4	1	2	8	7	7	4	7	5	8
5	1	1	2	4	5	3	5	5	5
6	1	0	1	3	5	2	4	5	5
7	1	0	6	4	5	2	5	6	8
8	1	0	2	3	4	2	5	7	7
9	1	0	6	7	8	0	5	5	4
10	1	2	5	6	4	1	4	5	5
11	1	1	2	2	1	3	4	5	5
12	1	1	2	4	6	2	3	5	7
13	1	2	3	5	4	2	3	4	2
14	1	0	2	4	6	2	2	6	3
15	1	2	1	2	2	0	4	6	6
16	1	0	6	7	6	2	5	7	7
17	1	1	6	7	7	1	3	6	7
18	1	0	8	5	6	2	4	6	4
19	1	5	8	8	7	2	4	4	6
20	1	0	8	7	7	1	6	7	7
21	1	0	5	4	7	2	4	5	5
22	1	0	5	5	6	0	4	2	2
23	1	2	5	3	5	2	4	5	5
24	1	3	6	6	7	6	8	7	6
25	1	1	4	5	5	2	4	5	5
26	1	1	3	4	6	2	4	5	5
27	1	1	5	6	4	2	4	5	5
28	1	0	4	5	5	2	3	4	4
29	2	6	8	6	7	6	7	7	5
30	2	2	8	8	6	1	5	5	7
31	2	1	3	4	1	2	4	5	5
32	2	5	7	8	8	3	4	5	6
33	2	3	7	6	8	3	7	7	8
34	2	0	2	6	5	2	4	5	5
35	2	2	7	7	5	1	4	5	3
36	2	4	7	8	6	4	5	6	5
37	2	1	3	5	5	2	5	6	6
38	2	2	3	4	5	1	3	3	2
39	2	3	7	8	8	5	7	8	7
40	2	3	6	8	6	1	5	5	7
41	2	4	6	7	8	1	6	8	8
42	2	1	4	7	4	1	5	7	7
43	2	5	7	6	8	7	8	8	8
44	2	3	5	5	5	1	3	5	4

45	2	2	0	1	3	0	0	5	4
46	2	3	8	6	7	3	6	7	7
47	2	1	2	3	5	1	3	5	5
48	2	4	5	7	7	3	7	7	7
49	2	1	3	8	3	1	6	7	5
50	2	6	8	7	8	4	8	8	8
51	2	4	8	7	6	4	7	6	7
52	2	5	6	6	6	3	4	6	6
53	2	2	5	5	4	2	4	4	4
54	2	6	8	5	8	5	8	7	7
55	2	4	5	7	5	1	3	7	6
56	2	4	4	5	6	1	2	5	4
57	2	3	3	4	5	2	3	4	6