LEARNER-COMPUTER TEXTUAL GLOSS INTERACTIONS FOR SECOND LANGUAGE VOCABULARY ACQUISITION

MOHAMAD ALI YUSUF

THESIS SUBMITTED IN FULFILMENT OF THE REQUIREMENT FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

FACULTY OF LANGUAGES AND LINGUISTICS
UNIVERSITY OF MALAYA
KUALA LUMPUR

2015

UNIVERSITI MALAYA

ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: Mohamad Ali Yusuf (I.C./Passport No: 600909-08-6379)

Registration/Matric No: THA 030020

Name of Degree: Doctor of Philosophy

Title of Project/Research Report/Dissertation/Thesis ('this work")

Learner-computer textual gloss interactions for second language vocabulary acquisition

Field of Study: Applied Linguistics and Computer Assisted Language Learning

I do solemnly and sincerely declare that:

(1) I am the sole author/writer of this Work;
(2) This Work is original;
(3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any except or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
(4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
(5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya ("UM"), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
(6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate's Signature Date: 2 January 2015

Subscribed and solemnly declared before

Witness's Signature Date: 2 January 2015
Name: Assoc. Prof. Dr. Su'ad Awab
Designation: Supervisor

Name: Dr. Tam Shu Sim
Designation: Co-Supervisor 2 January 2015

ABSTRACT

Working from an Interactionist approach to Second Language Acquisition (SLA), this study further investigated the potential of Gass' (1997) Input-Interactionist model for SLA in a Computer Assisted Language Learning (CALL) environment. The model is used as it is described as the most comprehensive and clear statement of the roles played by input and interaction in L2 acquisition currently available. The amalgamation of different hypotheses such as the Input Hypothesis, Interaction Hypothesis, Noticing and the Comprehensible Output Hypotheses form the framework of this model. This Input-Interactionist model was operationalised through the learners' interactions with a textual computer gloss (modified input). The output for SLA in this study is measured by vocabulary knowledge development in a series of vocabulary tests. Additionally, the study attempted to find out if the students' language proficiency affects the interactions and the outcome. The modified input which formed the glosses were created both at word and sentence levels and in the students' L1 (*Bahasa Melayu*) and L2 (English). A total of 99 students in a Malaysian public university accessed a story online where the unfamiliar words in the text were glossed in the students' L1 and L2 with meanings provided at word and sentence levels. Using ANOVA, the findings indicated that both word and sentence type of glosses in the students' L1 and L2 can develop the students' vocabulary knowledge. It was also seen that the students in the word *Bahasa Melayu* and sentence *Bahasa Melayu* interacted with the glosses frequently. It appeared that mid proficiency students interacted the most on the glossed words and the least was the students with high proficiency. This revealed that proficiency levels do influence learner-computer interactions to a certain extent. Data also revealed that sentence and *Bahasa Melayu* type of language aided the low and mid proficiency students while high proficiency students benefitted from their interactions with sentence, English glosses in

the short term. In sustaining vocabulary knowledge, mixed results were obtained. On the whole, the Input-Interaction model can, to a certain limit, be mapped onto a CALL environment on two accounts. Firstly, noticing appeared to be a feasible feature in CALL that can aid language acquisition. Secondly, interaction in CALL as in glosses may also be applied in a CALL context; however there appeared to be limitations when compared to interactions in a face-to-face SLA situation. There appears to be a need for richer and engaging type of interactions that may benefit the learners with multiple exposures to different types of glosses. In other words, the interactions with the computer-aided text gloss have to be more face-to-face like with the provision of negative feedback and modified output to realize the benefits of the Input-Interaction model in a computer context. Among the limitations to the research were the study was not able to distinguish between clicking and interaction. In addition, the small number of words which were glossed probably displayed less definite patterns of learner-computer interactions with the glosses.

ABSTRAK

Sebagai pembukaan, dinyatakan bahawa kajian ini dirangka dari pendekatan *Interactionist*. Seterusnya, kajian ini meneroka potensi model Input- Interaksi yang telah direka oleh Gass (1997) untuk pemerolehan bahasa kedua (SLA) untuk digunakan dalam persekitaran pembelajaran dengan komputer (CALL). Sehingga kini, model ini digambarkan sebagai satu kenyataan yang paling komprehensif dan jelas mengenai peranan yang dimainkan oleh input dan interaksi dalam pembelajaran bahasa kedua (L2). Penyatuan hipotesis yang berbeza seperti Hipotesis Input, Hipotesis Interaksi , *Noticing* dan Hipotesis *Comprehensible Output* membentuk kerangka model ini. Dalam kajian ini, model *Input-Interactionist* telah dilaksanakan melalui interaksi pelajar dengan komputer gloss teks (input diubahsuai). *Output* SLA dalam kajian ini diukur dengan pembangunan pengetahuan perbendaharaan kata dalam satu siri ujian perbendaharaan kata. Selain itu, kajian ini cuba untuk mengetahui jika penguasaan bahasa pelajar memberi kesan kepada interaksi dan hasilnya. Input diubahsuai yang membentuk glos telah diwujudkan pada dua tahap iaitu perkataan dan ayat didalam bahasa *Bahasa Melayu* (L1) dan Bahasa Inggeris (L2). Menggunakan ANOVA, dapatan menunjukkan bahawa pelajar-pelajar dalam tahap perkataan *Bahasa Melayu* dan ayat *Bahasa Melayu* berinteraksi dengan glos dengan kekerapan yang tinggi. Kelihatan juga bahawa pelajar bertahap penguasaan bahasa pertengahan yang paling banyak berinteraksi dengan gloss sementara pelajar kemahiran tinggi berinteraksi dengan gloss yang paling kurang. Ini menunjukkan bahawa tahap penguasaan bahasa mempengaruhi interaksi pelajar –komputer. Data juga menunjukkan bahawa gloss pada tahap ayat dan *Bahasa Melayu* membantu pelajar bertahap rendah dan sederhana, manakala pelajar berkebolehan tinggi bermanfaat daripada interaksi tahap ayat dan glos dalam Bahasa Inggeris dalam jangka pendek. Dalam mengekalkan pengetahuan perbendaharaan kata pada jangka masa lama, keputusan bercampur-campur diperolehi.

Secara keseluruhannya, model Input- Interaksi boleh, kepada had tertentu, dipetakan ke persekitaran CALL dalam dua keadaan. Pertama, *noticing* boleh menjadi satu ciri yang dilaksanakan dalam CALL untuk membantu pemerolehan bahasa. Kedua, interaksi dalam CALL seperti dalam glos juga boleh digunakan dalam konteks CALL, namun terdapat batasan-batasan berbanding interaksi dalam keadaan SLA muka-ke-muka. Nampaknya, ada keperluan untuk interaksi komputer yang pelbagai dan menarik untuk memberi manfaat kepada pelajar dalam pembelajaran bahasa kedua. Dalam erti kata lain, interaksi dengan gloss komputer perlu lebih ciri-ciri muka-ke-muka seperti peruntukan maklum balas negatif dan *output* untuk merealisasikan manfaat interaksi pelajar-komputer. Antara batasan kajian ini ialah ia tidak dapat membezakan antara klik dan interaksi. Di samping itu, bilangan kecil kata-kata yang telah digloskan mungkin tidak mepaparkan corak interaksi pelajar - komputer dengan jelas.

*glos – penjelasan makna perkataan yang ringkas

ACKNOWLEDGEMENTS

I thank the Al-Mighty for giving me the strength and perseverance to complete this PhD. It has been a long endeavour**,** with many hiccups, twists and turns along the way. But holding close to my late mother's principle that one has to end what one begins, I pushed on.

As in any long-term effort, this PhD experience has seen me receiving support from many quarters. Firstly, my appreciation goes to my supervisors, Associate Professor Dr. Su'ad Awab and Dr. Tam Shu Sim. To Dr. Su'ad thank you for the support and encouragement in all these years. To Dr. Tam, my deepest gratitude for letting me see the light in my work through the long hours of discussions.

My appreciation also to my friends Kevin, Izzamir, Dr. Arik, Dr. Lim Peck Choo, Ray, Chuah and Puteri for the cheers each time a milestone was achieved while working on this thesis. To my family, especially my brother Mydin and sister Rahimah, thank you for your patience.

I thank UiTM for giving me this opportunity to pursue my studies. Also to my colleagues in UiTM Perak for the cooperation and help extended to me while I worked on the thesis. Finally, I am grateful to the students of UiTM Perak who were willing to take part in this research.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF TERMS AND ABBREVIATIONS

BM	*Bahasa Melayu*
CALL	Computer-assisted language learning
EN	English
ESL	English as a second language
Gloss	Definitions or explanation of difficult words
L1	The first language. In this study, the L1 is *Bahasa Melayu*.
L2	The Second language. In this study, the L2 is the English language
Look-up Behaviour	It is defined as how the learners use the glosses which are provided in the study
NS	Native Speaker
NNS	Non-native speakers of English e.g. second language learners
Online text	The text "A Scary Night" uploaded onto the Internet.
Proficiency	It is evaluated on the learners' UiTM English language programme examination scores as well as their Sijil Pelajaran (SPM) English results. Based on these evaluations, the study has three levels of proficiency: low, intermediate and high.
SLA	Second language acquisition
Student/Learner	A person who is learning another language apart from his/her first language. The terms are used interchangeably in the thesis.
SBM	Sentence *Bahasa Melayu*
SEN	Sentence English
WBM	Word *Bahasa Melayu*
WEN	Word English

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Introduction

This first chapter introduces the background and rationale for the study. In addition, a brief description of the framework of the study is provided. Also, the reader finds the purpose of the study, its research questions, the significance and, the assumptions and limitations of the study in this initial chapter.

1.2 Background and rationale for the study

In Second Language Acquisition (SLA), input and interaction are two notions which are prominent. Broadly, input refers to the language which is addressed to the learners. It is a fact that input is essential for language learning (Gass, 1997; Gass & Mackey, 2006). Input that is provided to the receiver be it a young child or language learner has to be comprehended by the receiver so as to render the input meaningful for language development. Hence, at times the input is modified and adjustments made to the discourse which has been commonly termed in SLA literature as caretaker speech (for young children) or foreigner talk (for non-native speakers). Further adjustments are made to the discourse by both parties engaging in what is known as interactions to enhance the comprehensibility of the input. It is in these interactions that SLA researchers (Gass & Mackey, 2007; Hatch, 1978; Long, 1996) who claim that such interactions are beneficial for language learning. Interactions may refer to any kind of two-way exchanges between people who are having a conversation or dialogue through the use of linguistic or non-linguistic means (Chapelle, 2003).

In addition, Ellis (1999) has taken the stand that interaction may benefit language learners through what he terms as intrapersonal and interpersonal interactions which can be interpreted as cognitive and social benefits of learning a language. Given the prominence of interaction in SLA literature, there is little doubt of its role in language learning. In fact, it is widely accepted that "there is a robust connection between interaction and learning" (Gass & Mackey, 2007, p. 176). These two notions of input and interaction have been studied extensively by SLA researchers to unlock their roles and effects in SLA. With that as an overview of input and interaction in SLA, it is stated early in this thesis, that this study is carried out within the Interactionist Approach, whilst examining the applicability of Gass' (1997) Input-Interaction model in a computer environment.

Apart from SLA, this study also falls under the ambit of Computer Assisted Language Learning (CALL). CALL is a wide and evolving field which covers a wide range of practices in using the computer in the teaching and learning of languages; in sum CALL is a diverse area with a lot of versatility (Youngs, Ducate & Arnold, 2011). That being the case, there are calls from researchers such as Chapelle (1997) who sees that CALL research should be aligned to SLA theories. The argument here is that the findings from such research can contribute directly to second language acquisition/learning.

It is acknowledged that the Internet's importance in our lives has become more significant, not only in information seeking, trade, entertainment, communication and education. According to the Malaysia' Communications and Multimedia Commission (2012), among the purposes of using the Internet is 88.3 % is to look for information, for education is 63.5% and reading is at 57.2 %. These figures underscore the interactions that occur between the computer and the user.

With the advent of more advanced, powerful computers and the Internet, interactions of this nature has emerged which allows different types of interaction between the user and computer through the computer platform. These interactions have been termed by Chapelle (2003) as learner-computer interactions.

One then asks, apart from obtaining information how much of the information is comprehensible to these Internet users where English is probably their second language. The Internet World Stats states that English is the most widely used language in the Internet with 536.6 million users in 2010. The incomprehensibility of the information may be attributed to many factors such as reading in a different language, difficulty of text, lack of background knowledge and linguistic constraints such as lack of vocabulary. On one hand, the lack of vocabulary can be a major obstacle for second language learners as researchers (Grabe, 1991; Haynes & Baker, 1993; Laufer, 1997; Read, 2004) have found out that vocabulary is the main factor which can impede or enhance comprehension. On the other hand, research done by Schmitt (2000) and Grabe (2004) has also demonstrated that learners acquire vocabulary incidentally through exposure to the language such as in reading. In other words, reading online can be an avenue for language learning, specifically vocabulary learning.

With that as a background, this study begins the line of inquiry in SLA and CALL on several fronts. Firstly, what aspects of input and interaction can be investigated in a CALL context, and secondly, what features of the computer and technology that can be harnessed to make vocabulary knowledge development possible, which may facilitate vocabulary learning. Thirdly, would the learners' proficiency in the language have a bearing on the use of the computer features and ultimately affect the vocabulary knowledge development process. The investigation is carried out by assuming that with vocabulary knowledge development it can ultimately pave the way for SLA and better comprehension of what learners read on the Internet.

In learner-computer interactions, one of its benefits is for the learner to obtain modified input from the computer. From this perspective, there has been extensive research on learner-computer interaction in the form of computer glosses featuring technology at its best with the use of multimedia glosses. Thus far the narrative on this has been positive – they are successful for language learning - yet there have been instances where learners and researchers have stated that more direct, simple and straight forward glosses should be in place (please see sections 2.3.1.1 & 2.3.1.2).With that in mind, this study is designed and framed within the Interactionist Approach as it investigates the applicability of the Gass' (1997) Input-Interaction model in a CALL set-up by using textual-only glosses for vocabulary knowledge development.

1.3 Input-Interaction model of SLA

Heeding that call for aligning CALL to SLA, it is best at the outset to state that this study is framed on Gass' (1997) the Input-Interaction model of Second Language Acquisition (SLA). This model by Gass encompasses five stages in the process of language acquisition starting from input where input is processed and undergoes conversion into output.

Ellis (2008) notes that Gass' model is a basic computational process in SLA in the sense that it describes the information processes that the input undergoes in the various stages such as apperception where the input is apperceived by the learner, leading to comprehended input. This is then converted into intake into the student's internal mechanism where it is processed further and reinforced into the student's interlanguage or what Gass' terms as integration where the linguistic information is developed or stored.

The final stage in the model is output where Gass' writes as "an overt manifestation of the whole process of language learning" (p. 7) within the model. In other words, it is at this stage that the student produces the language either orally or in written format. Within this process are mediating variables between the different stages; among them are frequency, attention, noticing and negotiation.

Gass' model of Input-Interaction is primarily based on an oral or face-to-face environment where input is provided to the student or non-native speaker (NNS) by a more able interlocutor or a native speaker (NS). Subsequently there is interaction between the two interlocutors where there will be modifications to the input as a result of feedback from both the NNS and NS.

To explain more, during interaction, there will be instances where the interlocutors negotiate for meaning because of the lack of understanding. From this negotiation, learners receive negative or positive feedback on their language which would in turn draw their attention to the deficiencies in their language, leading them to notice the gaps between their own language and the target language.

Gass and Mackey (2006) write that conversational interaction together with input is the focus of the Interaction approach to understand SLA. They continue to say that it is the interactions between learners and native speakers (NNS/NS) or among learners (NNS/NNS) of the language being learned is where "language is negotiated or feedback is provided" (p. 6) which can lead to enhance comprehension and perhaps language learning. In simple terms, it is envisaged that these modifications in the input and interactions would result in language learning.

With that as a framework, this study maps Gass' model onto a CALL environment, where the focus is to examine the look-up behaviour of the learners with computer textual glosses, that is, the use of the different types of glosses and their effects on vocabulary development among students with different language proficiency

levels. The flexibility of CALL is its ability to provide enhanced input which can be varied by accommodating individualised learning. The next section gives a greater account of input and interaction in the SLA process before interaction in CALL is discussed.

1.3.1 Input in SLA

Language input is necessary for a language to be learnt. Gass and Mackey (2007) define it as "language that a learner is exposed to (i.e. from reading or listening)". They continue to emphasize that it is an "essential component for learning as it provides the crucial evidence from which learners can form linguistic hypotheses" (p. 177).

SLA theories do in one way or another account for input as part of language learning. This common ground on input stops here however as the nature and its role in language learning is disputed by the different SLA theories. Among some of them, for instance, early Behaviourist Theory laid out the environment was important for learning and target language input was seen as a stimulus for formation of habits. On the other hand, the Monitor Theory posits that language can be acquired if learners receive comprehensible input, while from the Sociocultural Theory puts forward that language learning is a socially mediated process involving input. Therefore it can be seen that the data or input for language learning is an important start to the complex process of language learning irrespective of varying theories. Since input is necessary for language development, the logical question emanating from this is how input is processed by learners and others to facilitate language development.

There are some ways of configuring input into simplified input, modified input, comprehended input and enhanced input for the benefit of the language learner which will be further discussed in Chapter 2. (Please see section 2.2.1.2 on Quality of Input). The next section deals with interaction and how input is shaped during interaction.

1.3.2. Interaction in SLA

In a face-to-face conversation, the interaction between a proficient speaker and a language learner offers conditions that can be potentially beneficial for a language to be learnt. This is because in an oral communication context, there are bound to be instances where there are gaps in the conversation, where the interlocutors face communication problems such as not being able to understand each other caused, by language problems. This will open up a conversational structure which would repair the conversation in order for that communication to proceed. Such exchanges can be termed as negotiation of meaning (Pica, 1994; Gass & Mackey, 2006). This negotiation of meaning can take on several forms, among them are confirmation checks, requesting for clarification, and comprehension checks, repetitions or recasts (Long, 1983).

As a result of these exchanges, modified input is provided by the more able speaker. Specifically, modified input, as the name suggests, is input which is adjusted accordingly to meet the interlocutors' understanding or even enhancing it. Modified input is the language that is addressed to learners in an oral type of interaction which is termed as foreigner talk in second language acquisition settings (Gass & Mackey, 2007). Gass and Mackey (2006) articulate that conversational interaction together with input is the focus in the Interaction Approach to understand SLA.

1.3.3. Interaction in CALL

What has been set forth at this point is that interaction can facilitate language learning. Further, it has been identified that interaction can be of three types. One is interpersonal between the interlocutors and the second is intrapersonal which refers to the interaction within the learner's mind, while the third is learner-computer interaction (Chapelle, 2003).

This process of interaction in CALL can benefit language development through three ways: obtaining enhanced input, interaction through the process of negotiation of meaning and directing attention to linguistic form (Chapelle, 2003). According to Chapelle, there can be three types of manifestation of interaction in CALL; one where there is negotiation of meaning over the computer networks in synchronous communication, for instance, the online chat. The second type of interaction is where learners obtain modified input through the computer. The third type of interaction refers to the internal processes which are found in the learner where he/she is focused on the linguistic form and there is cognitive processing of input.

Chapelle (2003) posits that the benefits of such interactions in CALL can be a way of getting better input to increase knowledge and understanding, and also to trigger processing of input. In the learner-computer type of interaction, one way a learner can obtain modified input is "by clicking on hypertext to receive modified input in the form of receiving help for comprehension or dictionary help" (p. 58). As the word "clicking" appears in Chapelle's statement on interactions in CALL, it is best on the outset to clarify that in this study the term clicking refers to the initial or the starting point towards interactions. In that sense, clicking is seen to be the precursor to interaction.

In relation to receiving help while reading online, one common feature in CALL is where online glosses with word meanings are provided to help in reading comprehension and vocabulary acquisition. The focus is on vocabulary as Ellis (1999) has pointed out, that negotiation of meaning is often quite explicitly on lexis (p. 52). Further evidence is provided by Smith (2004) who claims that "interactionist research clearly suggests that learners tend to negotiate around lexical rather than grammatical features" (p. 370).

According to Chapelle (2003) when learners received modified input, two critical issues arise. One is the quality of the modified input and the other is the extent to which learners engage in the interactions. The quality of the modified input would be types of modification which also happen in the oral face-to-face situation. Among the types of modification are instances where there are simplification, repetition, clarification or comprehension checks. The extent of the interactions would refer to how engaged the learner is in the interactions to potentially benefit from them. With regards to the presentation of the modified input, technology has an array of forms such as using text, images and multimedia characteristics in the presentation of modified input in the form of glosses.

1.4. CALL and gloss

Looking at the literature of the use of glosses and computers, it can be deduced that research in this area is not new. It can be said that research in the area has stretched over the last two decades. In fact, research in the use of gloss and computer has covered much ground, more so in recent times with the use of multimedia type of glosses for comprehension and language learning. The area of research has been diverse with studies with various theories backing them such as SLA or Multimedia theories.

This strand of research has also looked into the use of such glosses for vocabulary acquisition. Read (2004) writes that glosses with the CALL context has so far "demonstrated that the provision of glosses can assist vocabulary learning from text, without interfering with the reading process" (p. 154).

1.4.1 Type of gloss

In the review of literature of the type of glosses for this study, it has been revealed that glosses are generally helpful for learning vocabulary and comprehension (please see section 2.3.1.1). There has been a lot of research was done on type of glosses. To begin with, Cumming, Cropp and Sussex (1994) carried out research comparing gloss formats. The formats were word or sentence definition on its own, plus a usage example. They found out that sentence definition was most preferred by the ESL students. Similarly, researchers (Hulstijn; 1993; Knight, 1994; Chun & Plass, 1996; Grace, 1998; Lomicka, 1998; Laufer & Hill, 2000; and Yoshii, 2006) indicated that sentence-level definitions of words are helpful for students. However, other researchers (Chun, 2001; Gettys, Imhof & Kautz, 2001; Hegelheimer, 1998) on textual gloss have found out that word definitions and word translations are helpful for students in vocabulary development. It seems that both forms of glosses are beneficial with no clear distinction between them. Therefore, this is one of the questions that this study hopes to answer; are both types of glosses equally beneficial or does word or sentence type of glosses has the edge over the other?

1.4.2. Language of gloss

The language of glosses is also a debatable area. For instance, researchers (Cheng & Good, 2009; Jacobs, Dufon & Hong, 1994; Laufer & Hill, 2000; Yoshii, 2006) found that there was no difference in the role of L1 or L2 glosses. At the other extreme, Hayden (1997) revealed that students hardly consulted other glossing options when there were L1 glosses available. Lomicka (1998) in her sample of university students studying French found out that students had a preference for L1 glosses in definitional-type of glosses. There were also researchers such as Davis and Lyman-Hager (1997), Nagata (1999) and Li (2010) who found out that using the students' L1 in the gloss is beneficial for the students. In contrast, Miyasako (2002) found out that L2 gloss groups outperformed the L1 gloss group on vocabulary learning.

From the above description, the results of the studies between L1 and L2 show that the language of glosses for vocabulary learning is still inconclusive. Ko (2012) surmised this scenario well by writing that more studies on language of glosses are required as important questions remain unanswered.

1.4.3 Gloss use and proficiency level of students

From the literature review of gloss use, some researchers have pointed out that the proficiency level of the students does play a role in their gloss use or look-up behaviour. Ercetin (2001) and Jacobs, et al. (1994) found that the effectiveness of glosses varies according to the students' language proficiency. They argued that glosses have a different impact on students with different proficiency levels. Li (2010) who studied 20 Chinese ESL students' vocabulary retention discovered that both L1 glosses and bilingual dictionaries were effective for students with lower proficiency levels.

Miyasako (2002) and Taylor (2010) claimed that L2 glosses were more effective for students with higher level of proficiency, while L1 glosses were effective for lower proficiency levels. In contrast, Yoshii and Flaitz (2002) found out that there were no significant differences between beginning and intermediate students in the rate of change between the immediate and delayed vocabulary test scores. An early study by Bland, Noblitt, Armington and Gay (1990) revealed that students with higher L2 proficiency preferred L2 glosses. Along the same lines, Ko (2005) pointed out that if the level of proficiency is high enough to understand the definitions in L2, then L2 glosses were more effective than L1 glosses.

Knight (1994), Yoshi (2006) and Abraham (2008) have examined gloss use behaviour and agreed that the language ability of the students affects the way they use the gloss. Recently, Yun (2011) who carried out a meta-analysis of 10 studies on gloss use on reading comprehension and vocabulary learning revealed that proficiency was a statistically significant variable that had made an impact on L2 vocabulary acquisition on beginning students.

From what has been written on type, language of gloss and proficiency level of language learners, it is clear that the whole area of glossing is intense but with no clear understanding of how these variables contribute to SLA.

1.5 The Issue

Looking at the description of the use of glosses and computers, it can be concluded that findings from research in this area have not been distinct. Glosses have been proven to be generally beneficial for learning vocabulary in an already well-researched area; nonetheless, there are still a few issues that need to be addressed in the

area of computer glosses. The issues are: type of glosses, language of glosses, proficiency of students and gloss use.

A closer review of this area shows that there are learners who prefer more straight-forward glosses without the glitter and fancy of multimedia characteristics. Students prefer more direct, straight forward and easier glosses to process. Furthermore, it has been said that multimedia glosses can have the deleterious effect on learning when students' cognitive abilities are overloaded when they are confronted with such glosses (Mayer, 2001). Another important aspect not to be neglected is that there is a certain threshold, after which technology ceases to be effective. Hence Gettys, Imhof and Kautz (2001) caution that there could be "overruns" of the real needs of users. They argue that the software should follow the principle of "economy of means" (p. 92) rather than extend the information accessible to the learner. In other words, even with all the technology available, one has to reduce the amount of information to the degree of minimum sufficiency for such help to be efficient.

Next are the factors of practicality and ease of developing glosses given the immense volume of online materials which can be glossed. For instance, software designers and classroom instructors would find it difficult to develop complicated glosses for reading passages in their work for computer applications and classroom practices. Suffice to state, it would be less complex and taxing if more straightforward and simpler glosses are designed and used. It would be more cost effective in terms of time and effort. Hence, the way forward would be to revert to basic provisions of glosses as in providing meaning at textual level by using definitions and translations. In fact, there is evidence to show that such glosses are preferred by students (Chun & Plass, 1996; Chun, 2001; Chun, 2006; Levy & Stockwell, 2006). It does appear research on more elementary and clear-cut textual glosses is needed.

Henceforth, the study examines the use of such glosses in vocabulary development by focusing on textual-only glosses; word and sentence in the students' L1 and L2.

Following this are the issues of language of glosses and the proficiency of the students in using glosses. It is seen that there has been no clear answer to this question of whether L1 or L2 is more effective as the language of glosses for language learning. To compound matters, is the uncertainty of how students with different language proficiency levels make use of glosses and benefit from them. Therefore, more research has to be undertaken before these issues are ironed out.

Finally, the Input-Interactionist model in SLA and its interlink of hypotheses are extended to a CALL environment to examine if it is a useable and efficient model for language learning in a different context.

1.6 The proposed study

After the fundamentals of the background and framework of the research have been described, this section explains briefly the nature of this study. Firstly, a text was carefully selected and uploaded on the Internet. Secondly, words which were deemed unfamiliar to the students were highlighted in a different colour from the rest of the text to create saliency. Attached to these words were the glosses which provided meaning at word and sentence levels in the students' L1 (*Bahasa Melayu*) and L2 (English). Therefore, modified input was provided to the students and these glosses can be accessed by clicking on the target words, triggering learner-computer interactions. The outcome of the interactions was vocabulary development of the students which was measured by vocabulary tests evaluating different types of vocabulary knowledge carried out at specific intervals in the study – pre, immediate and delayed.

1.6.1 Purpose of the study

The purpose of this study is to investigate whether learner-computer interactions with a text-only gloss in the students' L1 and L2 by learners with different language proficiency levels, can facilitate vocabulary knowledge development. Theoretically, it seeks to find out if the Input-Interactionist model in an oral context can be mapped onto a CALL environment. From the above purpose, the following research questions are formulated:

1.6.2 Research questions

1. What is the clicking behaviour of the learners in (a) the different gloss conditions and (b) of different proficiency levels?
2. (a) Which specific type of modified input, that is, word or sentence and in learners' L1 or L2 in the interactions facilitates (i) perceived vocabulary knowledge (ii) productive vocabulary knowledge and (iii) receptive vocabulary knowledge of learners with different language proficiency levels?

   (b) Is the knowledge maintained over time?
3. Is language proficiency of the learners a factor in determining how the learners interact with the glosses and subsequent effect on vocabulary knowledge?

1.6.3 The methodology

The study uses a quasi-experimental pre-test/post test research design. It involves 99 first semester university students.

The students, based on their English subject grade in the national level examination *Sijil Pelajaran Malayia* (SPM), are then randomly assigned into different gloss configuration conditions. (Please see Appendix A for research procedure).

The students then read an online text "A Scary Night" (as in Appendix C) where the target words have been glossed in a different colour from the rest of the text to create an enhanced text that provides saliency. Students then clicked on the target words to access the glosses. One important source of data for the study is the interaction or process data (Chapelle, 2003) which is obtained from the tracking device embedded in the online reading text. The tracking data records the number of clicks made by the learners as well as the type and language of the glosses interacted.

Besides this, the study measures the vocabulary knowledge development of the target words by the learners using different kinds of vocabulary tests. One test measures the existing knowledge by measuring the students' perceived vocabulary knowledge of the target words at three specific points in the study. One before the treatment, that is, before the use of the glosses, immediately after the gloss use and a delayed test after three weeks of the experiment. Two other types of tests are also administered on the students. These tests measure the students' receptive and productive knowledge of the target words immediately after gloss use and later after three weeks. The purpose of these tests is to see if the vocabulary knowledge has been sustained or declined over time.

1.6.4 Significance of the study

The study is significant as it maps the Input-Interaction perspective for SLA in a CALL environment, that is, from a principally oral communication domain to a CALL one. The study hopes to contribute to the input-interaction perspective in CALL by identifying what kinds of learner-computer interactions can aid SLA.

The study would also benefit teachers who want to introduce gloss in their teaching of vocabulary through gloss use built within an online context. Teachers and software designers too would find the results of the study helpful in terms of what type of glosses to provide, for example, in L1 or L2 and in what mode, word or sentence level meanings.

Together, teachers, software and web-designers are able to build on this research to help them prepare the most suitable type of gloss for the different needs and language abilities of students. It will also provide insight into how the different types of textual glosses affect learners with different levels of proficiency.

At its best, this information can further add value to existing body of knowledge on glossing by providing this valuable input on the role of proficiency in gloss use. Perhaps more importantly, learners themselves can develop vocabulary knowledge or even learn vocabulary on their own while reading online if aided by such glosses.

1.6.5 Assumptions of the study

The first assumption that is made in this study is that the clicks made by the students signal the start of the learner-computer interactions which in this case would be the interactions with the glosses. Therefore, the clicks act as precursors or as the initial step to the actual interactions with the glosses.

Next, as established in SLA, interaction can occur between learner, environment and within the learners' mind as well. The former two variables are more observable and easier to document for SLA as compared to what goes in the learners' mind, that in their internal mechanism. Therefore, in this study as in most other Interaction studies, it is assumed that some kind of mental activity for processing language is going on. What is recorded however is the clicking behaviour of the learners.

It is assumed that the clicking by the students signal the start of the interaction with the glosses and to a certain extent reflects the inner workings the learners' mind. In other words, clicking starts the interactions with the glosses and it probably indicates that learners are thinking about their interactions. The researcher admits that this study is unable to investigate this cognitive element that occurs within the learners.

The following assumption is related to SLA. While studies in SLA focus on learning and acquisition, this study examines knowledge, specifically vocabulary knowledge. It assumes then that vocabulary knowledge forms part of the greater picture of SLA. Hence, the study does not delve into the issue of learning and acquisition as the focus of this study is to examine input and interactions and how they affect vocabulary development. Moreover, these two terms are difficult to define even by the scholarly community.

Ortega (2009) writes that "it is impossible to investigate these constructs and the related prediction" (p. 135). She continues to state that "in contemporary SLA terminology no such distinction is typically upheld" (p. 5). Hence, the terms "learning" and "acquisition" are used interchangeably in this thesis to reflect vocabulary development.

Another assumption is made on the students' computer literacy. It is assumed that most students have basic computer and Internet skills, that is, in accessing online reading materials and clicking on linked words.

1.6.6 Limitations of the study

This research has a few limitations. Firstly, the study is unable to make the distinction between clicking and interaction. It is presupposed that the clicking acts as a start to the interactions with the glosses.

Perhaps, a tracking time taken for a click and interaction may act as a delineator between the two notions. However, before that can occur, the issue of how much time constitutes a click or interaction has to be explored further.

Secondly, although the low number of words in the text although has some research advantages, it may have contributed to the inability of the study to reveal more definite patterns of interactions with the glosses. Also, the small number of words can also affect the sensitivity of the tests to capture the vocabulary knowledge development of the students.

Thirdly, the study had only utilized students from one institution, comprising only *Bumiputera* students. Therefore, the findings of this study cannot be generalized to a larger population.

1.7 Conclusion

This first chapter provides the background and justification for the study. Besides these, the chapter outlines the methodology, significance, assumptions and limitations of the study.

The study extends the applicability of the Input-Interaction model from a conventional face-to-face situation to a CALL context. By doing so it hopes to inform other researchers and teachers on the practice of using computers in language learning within an Interactionist approach in the SLA framework. Eventually, it is hoped that the findings will be able to help all the stakeholders in the Interactionist Approach, encompassing input and interaction to advance the teaching and learning of languages.

CHAPTER 2

REVIEW OF LITERATURE

2.1 Introduction

This chapter reviews the current and relevant literature that will form the theoretical framework for this study. Working from a SLA perspective, this chapter will discuss the Input-Interaction model as well as the composites of noticing, interaction and comprehensible output to understand and develop the theoretical underpinning of this study. The subsequent section of the literature traces the development of Computer Assisted Language Learning (CALL) with perspectives on language development through computer-learner interaction, specifically with computer-aided glosses. As the study implicates a CALL reading situation, there will be a brief discussion on reading. Aspects of vocabulary knowledge and measurement wind up this chapter.

2.2 The Input-Interaction model

In his discussion on input and interaction in SLA, Ellis (2008) distinguishes two approaches to studying the roles of input and interaction in SLA. He identifies them as a computational model of L2 acquisition or sociocultural theory. Under the first approach, a number of hypotheses have emerged such as the Frequency Hypothesis (N. Ellis, 2002), the Input Hypothesis (Krashen, 1985), the Comprehensible Output Hypothesis (Swain, 1995), the Interaction Hypothesis (Long, 1983, 1996) and the Noticing Hypothesis (Schmidt, 1990, 1995).

According to Ellis, Gass (1997) modulated these hypotheses into a model of L2 acquisition to explain the complex process of L2 acquisition. According to Van Patten and Williams (2007), a model describes process or sets of processes of a phenomenon. In addition, and more importantly, "the model may also show how different components of a phenomenon interact" (p. 5). In this literature review, some of these hypotheses will be discussed and the section on theoretical framework later will show how these hypotheses inter-relate to underpin this study. Before this section discusses more on the Input-Interaction model, it is briefly and simplistically stated here that the second approach to SLA as put forth by Ellis (2008) is the sociocultural approach based on Vygotsky's theory of mediated learning where social interaction is important for language learning.

In the Input-Interaction model proposed by Gass (1997), there are five different components that interact in the process of language acquisition. They are apperceived input, comprehended input, intake, integration and finally, output. Within these components or levels there are factors such as frequency and attention mediating between one level and the other (Gass, 1997; Wesche, 1994). The process begins with input that is given to the learner. This fact is undisputed in the SLA that input is necessary for the process of language acquisition to start (Gass, 1997). However, Gass also stresses that not all input is utilized by the learner for language learning. This may vary depending on whether the input is modified, controlled or limited (Gass, 1997)

At the next stage of the model, input is apperceived. In simple terms, this refers to the situation where the learner realises that there is something to be learnt from the input. In other words, there is a gap between what the learners know and what the new items to be learnt are. Gass explains that apperception is an internal cognitive process that primes the input for further analysis (1997).

It is here that the mediating factor of 'noticing' and its related concepts of attention, awareness and consciousness take place. The other mediating factors are frequency, prior knowledge, and affective factors. Frequency is embodied in the Frequency Hypothesis that states that learning is aided by the frequency of the linguistics items which appear in the input. Prior knowledge and affective factors, on the other hand, are in the domain of the individual learner and not in the linguistics environment.

At the next level is comprehended input where parts of the apperceived input are comprehended. Gass (1997) differentiates between comprehensible and comprehended input. The difference is that comprehensible input is controlled by the speaker, while comprehended input focuses on the hearer or learner. Comprehension of the input can take on several forms such as semantic information to deeper structural analysis.

From this level, the learner may progress to intake which is where linguistic material is processed. It is at this stage that Gass sees interaction interplay with the learner's innate knowledge of the language which can cause intake. Clearly, this is a psycholinguistic process that takes place within the learner's mind (Gass, 1997).

After intake, there are two possible outcomes in the next level which is integration. Firstly, there could be interlanguage development or there is something learnt or secondly, linguistic knowledge could be stored for future use. What this means is that there is understanding of linguistic material but the learner is unsure of how to integrate this new knowledge into his existing knowledge, hence the knowledge is stored until there is further opportunity to integrate the knowledge. Gass also points out that "storage component is more likely for vocabulary and smaller chunks of language than for large syntactic strings" (1997, p. 7).

The last level in the model is the output. It is at this stage that the earlier processes in the model culminate. Here the learners produce the language that have been integrated or conversely, it can restart the whole process of acquisition when

feedback is received on performance where after learners adjust their hypothesis of the language accordingly. It is at this stage, that Swain's Comprehensible Output Hypothesis is in operation as output where the learner produces the language learnt. This stage can serve as links back to comprehended input after the process of negotiation where there is negotiation as adjustments are made to the language which was earlier produced. Further, it can also serve as feedback into the intake component (Gass, 1997).

Drawing on Gass (1997) and Ellis' (2008) description of the Input-Interaction model, a simple graphical representation of the input-interaction process together with the interplay of hypotheses in SLA is given below.

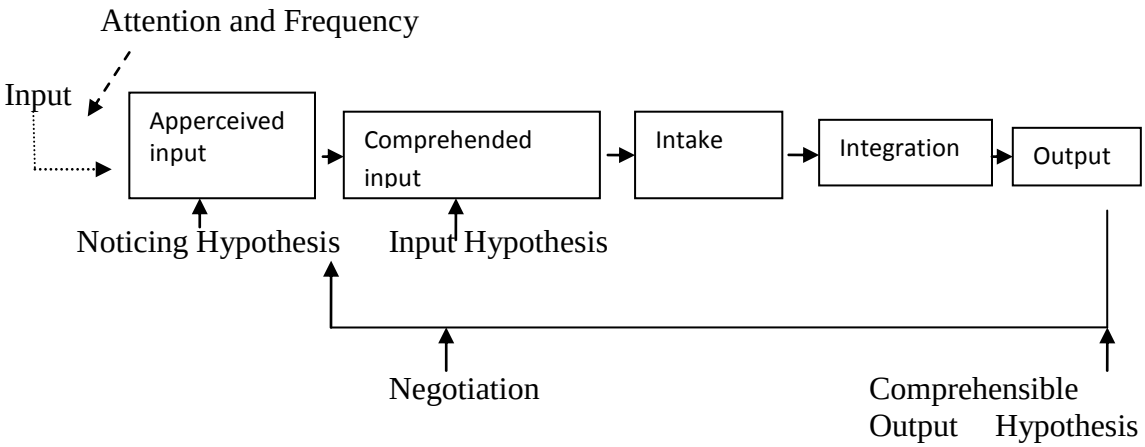


Figure 2.1: The Input-Interaction model and the interplay of hypotheses in SLA

The discussion so far has outlined the whole spectrum of processes in language learning from the Input-Interaction model. It can be seen that it is input that ignites the impetus for the language learning process. Given its nature as a trigger in the Input-Interaction model, it is discussed next.

2.2.1 Input in SLA

Essentially, input refers to the language which is addressed to the learners. It is a fact that input is essential for language learning (Gass, 1997; Ortega, 2009). A more comprehensive definition of input is given by Smith (1993) as "potentially processible language data which are made available, by chance or by design, to the language learner" (p. 167).

This definition by Smith is more exacting as it incorporates the elements of process as in 'processible language data' which is provided to the learners either by 'chance or design'. This means that input is whatever language that the learner is exposed to, either in instructed or non-instructed contexts. More importantly, this input has to be processed by the learner for it to be used successfully in language learning. What matters is the type and amount of input necessary for language development. As input is necessary for language development, the logical question emanating from this is how input is manipulated by learners and others to facilitate language development.

Among the possible ways of configuring input are simplified input, modified input, interactionally modified input and modified output. Modified output is also considered input as it restarts the whole process of communication as input emerging from the modified output (Gass & Selinker, 1994).

2.2.1.1 The Input Hypothesis

When input is discussed it is necessary to examine Krashen's Input Hypothesis as part of his Monitor Model in language acquisition. Krashen puts forth the case for comprehensible input in the form of the Input Hypothesis.

This hypothesis states that input is necessary for language acquisition and it has to be a little ahead of the learners. Krashen defines a learner's current state of knowledge as i and anything above that is known as + 1. Krashen continues to say that the + 1 stage is crucial for language acquisition to occur (Krashen, 1985). Anything that is below the 1 stage or any structures which are way ahead beyond the learners' current knowledge will not offer any benefits to language acquisition (Gass & Selinker, 2008; Krashen, 1985). Krashen claims that comprehensible input is not just necessary for SLA, but it is the sufficient condition to realize SLA.

Krashen's stand on input has received criticism on several fronts. Firstly, there is no clear definition of the levels of knowledge. Gass and Selinker (2008) question how would language instructors know the level of their input for the learners to benefit from the i + 1 formula. Furthermore, it would also be almost impossible to identify the different input for different learners who have various levels of language knowledge, meaning, the different *i* 's. Secondly, is the quantity of the input. Krashen states that there has to be adequate quantity of suitable input. There was no mention of the amount of suitable input for the learners. How would language instructors know the quantity which is sufficient for the learners? These learners could be at varying levels of language development and preparedness to acquire new forms of the language (Gass & Selinker, 2008).

Another criticism of this hypothesis is that Krashen looks at comprehension as a dichotomy; either something is understood or it is not. Gass' (1997) view on this is that comprehension is made up of different levels of comprehension such as semantics to structural analyses and it is not dichotomous as claimed by Krashen. While Krashen argues for the need for comprehensible input which can lead to language development, another directly contrasting view is the incomprehensible input which is the key to language learning.

White (1987) proposes the Incomprehensible Input Hypothesis where she contends that it can give rise to L2 acquisition. The claim was that it is the incomprehensibility of the input that would trigger the language development of the learner. The incomprehensibility of the message would force the interlocutors to navigate around the input by providing negative feedback and making necessary adjustments to arrive at meaning. Thus, incomprehensible input is said to be vital for SLA. Although, there are different stands on input and its role in SLA, there is no doubt of its centrality in the process of SLA. It can be seen that input is an important issue that needs to be treated in depth. Therefore, the next part of the review discusses the issue of the quality of the input.

2.2.1.2 Quality of input

In the past, there has been research on the quality of input. According to Van Patten (as cited in Lai & Zhao, 2005), input can take on various forms such as comprehensible input, simplified input, modified input and enhanced input. Early research was on how to manipulate input in order to facilitate learner's comprehension. This was proposed by Krashen in 1985. From here, researchers were interested to find out what aspects of the input were made comprehensible to the learner, leading to what is known as simplified input.

In simplified input, linguistic adjustments such as shorter sentences or the use of high frequency vocabulary were used. Later, Long (1983) introduced the concept of modified interaction which is resulted from negotiated interaction in which interlocutors modified their input to repair oral communication breakdown.

Next is enhanced input which can be defined in several ways. For instance, in conversational structures one of the ways is in the form of stress on certain words or phrases. In a CALL context, however, enhanced input can be in the form of marked input where input can be made salient by highlighting the structures that the learners are supposed to attend to while they are reading the text (Chapelle, 2003).

Additionally, input can also be modified by the interlocutors who take part in a conversation. The input is modified to make it more understandable to the less able interlocutor or the language learner. What matters is that input is an important initial step to language learning as it is from this input that learners begin to form linguistic hypotheses (Gass & Mackey, 2007).

It is acknowledged that input is necessary for language learning however as Ortega (2009) asserts, "input is undoubtedly necessary but it cannot be sufficient" (p. 60). In other words, there has to be other ways to manipulate input so that it can be better processed by the learners to facilitate SLA. This assertion brings forth another aspect of examining the notion of interaction.

2.2.2 Interaction in SLA

Research on interaction began with a focus on interaction in first language acquisition. Many studies in the area concentrated on mother-child interactions. These interactions provided conditions and input that could lead to language development in children. As the setting of language acquisition in these contexts was within social interaction, it was befitting that language development must be studied in a more interactional context.

Building on findings and observations of interactions in the first language acquisition, interaction has since been researched in L2. Such research notably focused on interactions between native speakers (NS) with non-native speakers (NNS). Research was primarily based on conversations between NS and NNS and highlighted some of the processes where both NS and NNS take part to repair or improve communication. Interlocutors modify and restructure their process of interaction to achieve mutual understanding. Among the processes are asking and giving clarification, repetition, rephrasing, providing elaboration, having comprehension checks to make meaning clearer; in sum, interactions to make the input more comprehensible or to repair breakdowns in communication.

In the field of SLA, interaction has been studied widely. Ortega (2009) documents that there have been two generations of studies on interaction alone. Each phase contributing actively to the realm of interaction for SLA. The first generation of studies focused on negotiation of meaning and comprehension. Notable researchers such as Pica (1994), Gass and Varonis (1994) and Loschky (1994) described the process of negotiation and identified factors which were related to interlocutors and tasks that generate interaction.

The second generation of studies on interaction extended the scope of interaction to include the link between interaction and acquisition. The studies on this were sparked off from Mackey's (1999) hallmark study which showed positive results of interaction and acquisition. Research from this phase spotlighted on three areas: (a) product-oriented designs that include pre and post-tests (b) the measurement of learning gains on particular forms and (c) the element of noticing was included. It can be deduced that interaction has been examined extensively to inform researchers of its role and contribution in SLA.

In the Interactionist perspective, interaction is between the learner's mental abilities and the linguistic environment. This perspective assumes that language learning takes place through social interaction. The Interactionists' view on second language acquisition has been that input is not only the necessary element for language acquisition to occur, but what is more important is the language or discourse that both learners and their interlocutors, who are proficient speakers, jointly construct.

Interactionists believe that interactions between a learner and a native speaker can lead to language acquisition when the learner engages in modified interactions with the native speakers through the negotiation of meaning. Nonetheless, the interactionists also do not discount the role of input, in particular, interactionally modified input which is embodied in Long's (1983) Interaction Hypothesis (IH).

2.2.2.1 The Interaction Hypothesis

It is important at the onset to clarify that the Interaction Hypothesis (IH) embodies some aspects of the Input Hypothesis (Krashen, 1982, 1985) and the Output Hypothesis (Swain, 1985, 1995, 2005). Therefore, when the IH is discussed, both the Input and Output hypotheses have to be examined too. The Input Hypothesis by Krashen has been discussed in the previous section while the Comprehensible Output Hypothesis is discussed later in this section.

The early beginnings of the Interaction Hypothesis (IH) can be traced back to early works in ethnomethodology which focused on how native speakers repair breakdowns in communication (Ellis, 1999; Hatch, 1978) provide useful understanding of how learners are able to learn a second language through the process of interacting in oral kinds of settings. Besides Hatch, there were many prominent researchers in the Interaction Hypothesis (IH) such as Pica (1994) and Gass (1997).

These researchers have studied the ways how native speakers and language learners interact in mostly conversations-type of contexts to aid language development.

Long (1983) in particular believes that although input is necessary for language acquisition, he further asserts that it is the modified interaction that is the necessary element for better language comprehension and possibly acquisition. This position differs from Krashen's stand that comprehensible input in itself is sufficient for language acquisition. Krashen's Input Hypothesis (1985) claims that learners learn a second language when they are able to access comprehensible input through processes when input is simplified or when learners use the context to make meaning clear. Long, on the other hand, focused on how input could be made more comprehensible through the interactions between language learners and their interlocutors to reach comprehension (Lightbown & Spada, 2006).

Long also had two versions of the IH, the initial one in 1983 and this was revised in 1996. At its inception, IH was based primarily on input and takes place in the context between a competent and less competent speaker of the language. The IH looked at how input was made comprehensible by the less competent speaker who provided feedback on his/her own lack of comprehension to the more competent interlocutor. The interactional modifications which are in fact changes made to the structure of an oral discourse undertaken by both the interlocutors to enhance understanding of the conversation. It is these modified interactions that assist in acquisition (Long, 1983).

Modified interaction may not necessarily involve linguistic simplification as it also can include elaboration, the provision of additional contextual clues, clarification requests by the learners or paraphrase or comprehension checks by the native speakers (Lightbown & Spada, 2006).

The IH was later revised by Long (1996) to include factors such as the discourse between the learners and the competent speakers including the role of feedback, and output from the language learner (Ellis, 2003). The revised version of IH now is more encompassing as it looks at input, modified interaction, negative feedback, learner attention and the subsequent output. It is also in the later version of the IH, that Long addresses how interactionally modified input helps in acquisition by making explicit the learners' internal mechanisms. One manifestation of the learner's mechanism is noticing, that is, when learners notice the linguistic forms in the input which cause a gap in their understanding of the input.

In discussing the IH, Ellis (1999) states that the "general claim of IH is that taking part in interpersonal oral interaction in which communication problems arises and are negotiated which then facilitates language acquisition" (p. 4). Ellis concludes that "The essence of an Interactionist perspective is that interaction, interpersonal and intrapersonal, plays a major part in creating the conditions in which language acquisition (first and second) can take place (p. 3)." Further, Saville-Troike (2006) takes up this point and concludes that IH can be surmised as when "L2 is acquired in a dynamic interplay of external and internal processes with interaction facilitating but not causing SLA" (p. 111).

At this juncture, it is imperative to make the distinction between interactional modifications and input modifications. Long (as cited in Ellis, 1999) defines interactional modifications as changes to the structure of a conversation to account for real or potential problem of comprehension, while input modifications refer to the native speakers' adjustments in conversation while addressing language learners. While input and interaction may point to successful language acquisition, it is apparent that there is a missing factor in the process of SLA.

Therefore, the sequence in the input-interaction cycle is inadequate if Swain's (1985) Comprehensible Output Hypothesis is not brought into the light of SLA.

2.2.2.2 The Comprehensible Output Hypothesis

In order to complete the input, interaction output sequence of language acquisition, this section discusses the issue of output. Although comprehensible input and interaction are plausible factors to language acquisition, there was reason by researchers to seek the missing link in the SLA process. This missing link could be the output from the learners themselves.

In 1985, Swain suggested that "output" was the missing factor and introduced the concept of comprehensible output constituting the Comprehensible Output Hypothesis. The point in this Hypothesis is that comprehensible output occurs when learners start to use the language items which have been previously learnt. Output can be in various forms, for instance, spoken or written. It refers to language which is produced by the learners, that is, when "learners are pushed in their production as a necessary part of making themselves understood" (Gass & Selinker, 2008, p. 326).

It is proposed by Swain that the hypothesis can provide three opportunities for the learners to develop their language. Firstly, to allow learners to have opportunities for contextualized meaningful use; secondly, it allows learners to test their language hypotheses; and lastly to make learners move ahead from semantic to syntactic processing of the target language.

Appealing as it may be, the role of comprehensible output in language acquisition is still largely unanswered and at best controversial Shehadeh (2002). Ellis (2008) in his opinion, suggests that output does contribute to language acquisition, what he contends is that it is not clear whether output helps learners to acquire new forms or only to automatize use of partially acquired forms.

It is evident that in the process of language learning, output is important but a mediating factor before output is produced, is the notion that the language items learnt will have to be first 'noticed'. This notion of noticing is discussed next.

2.2.2.3 The Noticing Hypothesis

Richard Schmidt (1990) proposed the Noticing Hypothesis (NH) which in its essence states that if there is to be any language development, the items have to be first noticed by the language learners. Schmidt documented his own second language learning experience of learning Portuguese. He discovered that the language that he noticed were of two forms: either they were brought to his attention or some other kind of experience had made them salient (Lightbown & Spada, 2006; Schmidt, 1990 ;).

It is said that the language learner must be aware of the linguistic input for it to be learnt. But before that can happen, Schmidt (1990) argues that attention is essential to learning. Attention and its related concept of noticing which forms part of the cognitive process is thought to be important in SLA. For Schmidt, "intake is that part of the input that the learner notices" (p.190). Schmidt also points out that all noticing is conscious. To elaborate, he continues that a learner in noticing something is conscious of the fact of noticing. In his discussion on noticing, Schmidt equates awareness to consciousness. Therefore, in his point of view noticing involves awareness and consciousness.

Along this same line on noticing, Gass explains that noticing "includes awareness, and awareness presupposes attention, and attention is central to any concept of noticing" (Gass, 1997, p. 8).

The centrality of attention is also acknowledged in learning by Gass and Mackey (2007). According to them attention is one of the mechanisms that "mediate between input and learning" (p. 18). Therefore, it is accepted that learning will take place when the learner is aware of the linguistic input for it to become internalized. They also remind researchers in the interaction perspective that attention and awareness form part of the interaction-L2 learning process.

In extending the notions of noticing and attention, Ellis (2008) makes a point that Schmidt distinguished three subsystems of attention based on the work on Tomlin and Villa (1994). The first is attention as 'alertness' which refers to learners' motivation and readiness to learn, and that noticing and acquisition do not hinge on the learner's intention, that is, involuntary noticing can also take place. Secondly, 'orientation' refers to the general focus of attention. It states that orientation may lie in the design of the task, that is, whether the focus of the task is on meaning or form. The third subsystem is 'detection'; it is what is registered in the mind that allows for further processing of information. It is this third subsystem which triggered a controversy of whether detection involves awareness on general attention to form or attention to a specific aspect of language. This question made Schmidt re-examine the Noticing Hypothesis and differentiated it into the strong or weak form. In its strong form, Schmidt's assertion is that there is no learning at all if input is not noticed and that attention has to be specifically directed. In its weak forms he contends that there is some form of learning as there is representation and storage of unattended stimuli in memory. In other words, there is some learning taking place without attention.

It has to be pointed out here that all the hypotheses in the Input-Interaction model of second language acquisition do acknowledge the role of consciousness in varying degrees. For instance, at the extreme end, the Input Hypothesis rejects the role of consciousness as Krashen claims that acquisition is a subconscious process, while Long in his IH makes the opposite claim that learners have to be conscious to benefit from negotiated interaction. In the same thread is Swain's Comprehensible Output Hypothesis which makes the point that consciousness is important for the learners to notice the gaps in their interlanguage and develop metalinguistic awareness.

It is important here to set the stage to elaborate how the elements of noticing and attention can be beneficial for SLA in a research context, for example, for learning vocabulary in a text. Nevertheless, before attention and noticing can occur, another element which should not be overlooked is saliency. It is believed that "saliency can affect acquisition" (Gass, 1997, p. 19). Saliency in SLA can include elements such as stress, word position and even frequency.

In vocabulary learning, the saliency of the target words may trigger learning. The target words can be made salient in several ways, for example, by making them more frequent in the text or by highlighting the words differently from the other words in the text which can also be termed as input enhancement. This is done in what is termed as "researcher-driven or externally-driven" (Gass, 1999, p. 321). This means, the onus of bringing the attention of the learners to the words is undertaken by the instructor or researcher.

A study done by deRidder (2002) on noticing found that when learners are presented with marked glossed words, they would be more willing to consult their meaning compared to if learners read a text without marked glosses. However, she discovered that consulting the word meanings did not have any gain in either vocabulary learning or reading comprehension.

The researcher is indeed cognizance of the difficulty of pinning noticing on linguistic data on its own. Hence, the need of documenting or recording the learners as they interact with the input as provided in the study by the students' interactions with the gloss. This in itself is also not direct evidence of noticing but it comes as close as possible to clarifying the construct noticing.

As opposed to this external noticing, there is another type of noticing. It is the learners themselves who notice the words. What is meant here is that learners notice that there is a gap in comprehensibility when they are reading the text and are faced with the unfamiliar words; hence it is called learner-driven or internally-driven. Gass (1999) emphasizes that both types of noticing are equally important for acquisition and have to be taken into account. These two forms of noticing are shown in the diagram below.

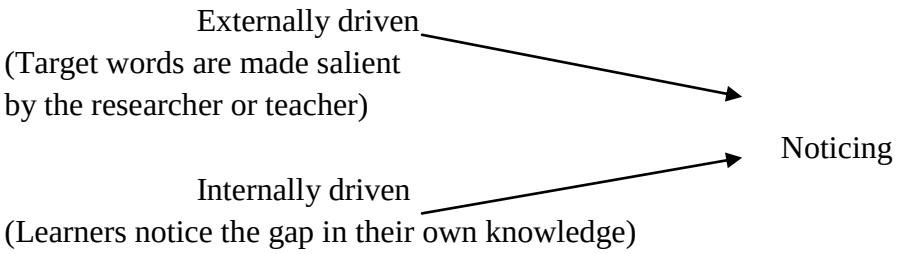


Figure 2.2: The two types of triggers which can induce noticing

It is through interaction with the text that attention is focused on a part of the language which has caused communication or comprehension problems. Gass and Selinker (2006) illustrate this when learners notice the deficiency of their between learner-language forms and target language forms. In this same vein of attention, is the fact that learners' requests for help also signal the instances of noticing (Hegelheimer & Chapelle, 2000).

This claim then puts it that it is through the process of getting help or feedback in the form of negotiation of meaning or interaction that the notion of noticing is heightened. This is supported by Gass (1997) who states that "the input-interaction view must take the position that noticing is crucial" (p. 100). In negotiation the learner is focusing on linguistic form, and that focus, or specific attention paid to linguistic form, is the first step toward grammar change

2.2.2.4 Negotiation of meaning

In an oral discourse between a competent speaker and the language learner, there are times when the learner fails to understand what is spoken and the learner may respond in two ways. Firstly is to pretend that he/she understands and secondly, the learner and the competent speaker engage in what is termed as interactional modifications to the discourse which constitute negotiation of meaning (Ellis, 1997).

Negotiation of meaning occurs in conversations where the L2 learner and the native speaker work together to overcome the communicative difficulties which are bound to happen with the learner's limited L2 ability (Ellis, 1985). Lightbown and Spada (2006) write that when communication becomes problematic, the interlocutors must negotiate for meaning. They maintain that it is this kind of negotiation that can create the opportunity for language development. Negotiation is seen as a vehicle that enables learners to obtain input which is comprehensible.

Parallel to this, Pica (1994), who worked on Krashen's (1985) premise that comprehension of meaning suffices to enable learners to access the forms and structures that encode that meaning, describes that negotiation facilitates comprehension of L2 by attracting learners' attention to form-meaning relationships. When there are modifications in the discourse, negotiation takes place and it initiates L2 comprehension

and this comprehension can lead to L2 learning and acquisition of L2 forms. In other words, the learner's comprehension of meaning can be the result of their access to L2 form. Pica states that "this twofold potential of negotiation - to assist L2 comprehension and draw attention to L2 form – affords it a more powerful role in L2 learning than has been claimed so far" (p. 508).

Pica (1994, 1996) details that negotiation helps comprehension of L2 input and draws the learners' attention to form-meaning relationships through the process of repetition, segmentation and rewording. The features that play a role in such negotiated learners' interaction are requests for message clarification and confirmation to make input comprehensible. The more competent interlocutor follows these up with some possible responses such as repetition, elaboration or simplification. These modifications can include at word, phrase or sentence levels. She states that it is common in these modifications to consist of checks for vocabulary which forms a significant factor of the negotiation process.

The point on vocabulary being the initiator of negotiation is continued with Smith (2004) who also finds that in a computer environment, learners tend to negotiate around lexical rather than grammatical features. This is in consonance with De la Fuente (2003) who writes that negotiation which is centred on lexical aspects of the language may be beneficial for L2 vocabulary acquisition. Hence, from these observations (De la Fuente, 2003; Pica, 1994; & Smith, 2004), it is collectively suggested that vocabulary is the common initiator to the negotiation process.

In a CALL environment, there can be similar instances of negotiation of meaning which are afforded by the computer, either by learner-learner where the computer acts as the medium of negotiation or more direct interfaces as in learner-computer interactions.

2.3 Interaction in CALL

Interaction serves as a good starting point where researchers in CALL can apply the tenets of the input-interaction theory to their field of study. In CALL, this interpersonal activity takes place over the computer network. It emphasizes the joint contributions of the linguistic and computer environment with the learner's internal mechanisms in language development. Learning results from interactions between the learner's mental abilities and the linguistic input and interaction provided by the computer.

As discussed in the earlier sections, interaction is a term used to refer to interpersonal activity that arises during oral communication (Ellis, 1999). In a CALL environment Chapelle (2003) states that one of the cornerstones of CALL is that it usually has some element of interactivity in it. Chapelle (2005) explains that she uses the term "interaction" as the superordinate concept that includes any type of two-way exchanges through the use of linguistic as well as non-linguistic exchanges" (p. 54). It is through the interactional acts that linguistic input can enter the learners' interlanguage system. The exchanges that take place have greater potential for language development than in activities where interaction does not occur. This process of interaction can benefit language development through three ways which are interaction through the process of negotiation of meaning, obtaining enhanced input, and directing attention to linguistic form (Chapelle, 2005). On top of these learners are also able to interact with other learners through the computers as in computer mediated communication (CMC) and other social interactive means.

At another level, interaction between learner and the computer is also provided where learners are able to interact with computers to obtain a host of materials, for example, information or news which is provided on the Internet.

In addition, the learner can obtain "enhances" or modified input from the computer. For example, when a learner uses the computer during reading or listening, and if she stops the input to ask for help in the form of meaning of difficult words, grammatical aid or even text transcripts can be paralleled to benefits of meaning negotiations in oral face-to-face communication instances. Recent research on CALL has found positive findings that support such learner-computer interactions, for example, in the acquisition of vocabulary (Plass, Chun, Mayer & Leutner, 1998; Lyman-Hager, 1996).

Besides these learner-computer interactions, there is another type of interaction in the form of intrapersonal interaction, that is the internal workings of the learners, or in short, the interactions that go on in the minds of the learners which are largely unobservable as compared to interactions between learners and learner-computer.

To summarise, in a CALL environment Chapelle's interpretation of interaction includes not only interpersonal and intrapersonal but also learner and computer. What can be summed up here is that these different interactions have their own benefits that can aid SLA as shown in Table 2.1.

Table 2.1: Benefits of three types of interaction

Basic types of interaction		**The Benefits of the interactions**
Inter	between people	Negotiation of meaning
	Between person and Computer	*Obtaining enhanced input*
Intra	within the person's mind	Attending to linguistic form

Note: Taken from: Chapelle (2003). *English Language Learning and Technology* (p. 56)

Chapelle (2003) notes that one of the CALL interpersonal interactions, is the "interaction between a person and the computer" (p. 55). In this kind of interaction, the benefit is to obtain enhanced input from the computer where the computer can provide useful and meaningful modified input to facilitate language acquisition.

In this particular aspect, there is no dearth of research as there are numerous studies in the form of providing glosses while learners read or listen to text using the computer. In such a situation a learner may interact with a computer to read an online text and is able to negotiate with it to obtain enhanced input from the computer.

2.3.1 Gloss and CALL

Gloss can be defined as providing information on important words through definitions or synonyms (Nation, 2001; Myong, 2005). Robb (1999) defines it as "glosses are many kinds of attempts to supply what is perceived to be deficient in a reader's procedural or declarative knowledge" (p. 96). Providing gloss to help in reading is not a new phenomenon.

According to Lomicka (1998), it dates back to the Middle Ages. Robb (1999) in his article "What's in a gloss?" draws some historical notes on the origins of gloss from lexicographer Werner Hullen who chronicles that early glosses were in the form of interlinear or marginal scribblings which were learner-generated. Medieval learners produced glosses when studying Latin. Glosses were used as teaching aids very much later on, followed by their "codification into word lists and then later dictionaries" (p. 94).

These days glosses are provided with the help of computer and are sometimes referred to as electronic glosses. Lenders (2008) categorizes such glosses into three types.

First is dictionary-type of glosses where they mirror information that is given by printed dictionaries, consisting of among others, phonetic script and usage example sentences. The second type is ready-made glosses where the glosses are tailor made for particular needs of learners and they provide meaning of a glossed word in a given context. Thirdly are special type of glosses which are often specific to the target works but provides extra information such as a task for the learner, for example, providing more than one possible translation of the glossed word in a multiple-choice format.

Lomicka (1998) expresses that this development triggered a lot of research on glosses which has remained largely unexamined until recently. Although this may be a positive forward step, she laments that it brought open the controversy and problem related to this area of providing gloss aid. The controversy and problems of gloss use are highlighted when the study looks at the different studies reviewed, especially in the area of computer gloss where different types of media are used, with different types of learners in various settings such as ESL, EFL, which may have single or dual objectives of enhancing reading comprehension and/or language learning. Furthermore, research into the area is carried out using different perspectives and theories, ranging from SLA, multimedia to vocabulary learning ones.

With such a conundrum of research on glosses, the findings from the research are diffused and it is difficult to streamline the findings to make them meaningful. Therefore, this study cautiously focuses on core and pertinent issues such as if textual glosses (word or sentence) can help students in developing their vocabulary knowledge. Next, the study is conducted within the realm of SLA, precisely drawing in Gass' (1997) Input-Interaction Model to explain the whole process of input and interaction of glosses for vocabulary learning.

Nation (2001) describes three main reasons why glosses are useful to learners. His first reason is that more learners are able to handle more difficult texts as a glossary of the unfamiliar words are provided. Secondly, with the aid of a gloss, learners will not guess the meanings of the words wrongly as the correct meanings are given. Thirdly, he affirms that glosses do not intrude into the learners' reading when compared to perhaps in using a dictionary.

As to the placement of glosses, the literature on gloss has also provided evidence that marginal gloss best aids the learners in reading and to some extent in vocabulary learning. This is also backed by Watanabe (1997) who voiced that "the main purpose of the glosses is to aid text comprehension, however, glosses are also used to promote vocabulary learning" (p. 289).

Another plus point for using gloss is provided by Ko (2005) who lists two advantages of using gloss in reading. Firstly, she lists that it helps in reading comprehension and vocabulary learning. Secondly, glossing does not interrupt the reading process. What is meant here is that since gloss provides definitions for low frequency words, readers do not have to stop reading to constantly look up their meanings. She goes on to explain that this will encourage interaction among three variables: the gloss, the reader and the text. This she says "may promote comprehension and retention of the content of the text" (p. 2).

Nagata (1999) in her study entitled "The Effectiveness of computer-assisted interactive glosses" neatly explains the success of glosses in four factors. Firstly, marginal glosses are easier to use than a dictionary. Secondly, glosses which are highlighted draw the reader's attention to them, hence giving credence to consciousness and input enhancement. Thirdly, glosses also connect words to meanings which allow readers to make the meaning-form connection as advocated in reading theories.

Lastly, it promotes depth of processing as readers use the gloss and target words which encourages the readers to perform lexical processing which may lead to vocabulary development and learning.

Other researchers, working in the realm of computer-aided glosses (Chun & Plass, 1996; Gettys, Imhof and Kautz, 2001; Lomicka, 1998; Lyman-Hager, 1997; Martinez-lage, 1997; Nagata, 1999 ;) studied the efficiency of glossing with positive results. Their studies have shown that L2 learners do get benefits from reading online glossed materials. Gettys, Imhof and Kautz (2001) also point out the advantages of using online glosses. They state, "… compared with traditional paper-pen-dictionary methods, online glosses enhance general comprehension, improve vocabulary retention and save learners' time and effort in reading L2 texts" (p. 91).

It is important to point out that with the use of computer gloss, learners are able to receive some form of improved input which involves two issues. One is the "quality of the modified input" and, secondly the "extent to which learners engage with the gloss which then opens up the possibility of learners deriving benefits from the interactions" (Chapelle, 2003, p. 59). It should also be noted that she recommends that for CALL to be useful for language acquisition, its effective use would include the following. One is to make key linguistic items salient by highlighting and providing opportunities for repetitions and modifications for particular forms. Secondly, it should support modified interaction by letting the learner have control when they need help.

2.3.1.1 Type of gloss

The next section of this chapter discusses how in this technology-age, diverse opportunities of utilizing the computer to provide gloss to readers.

Technology, in particular hypertext and computer multimedia capabilities, have enabled glosses to be presented in different media to the learners to aid in online reading comprehension or provide means for language learning.

When the type of gloss is discussed, it takes on two strands of research. One is textual-only gloss and the other is multimedia gloss. Early research on the type of gloss was carried out by many researchers. Among them are Cumming, Cropp and Sussex (1994) who carried out a research comparing word definitions formats. The textual formats were word or sentence definition on its own, plus a usage example. They found out that sentence definition was most preferred by the ESL learners, more so if there is an additional usage example.

Studies done by researchers such as Chun (2001); Gettys, Imhof and Kautz, (2001); Hegelheimer (1998) and Foroogh Azari (2012) on textual gloss have found out that word definitions and word translations are helpful for learners in vocabulary development. In this same thread of textual glosses, other studies (Hulstijn, 1993; Knight, 1994; Chun & Plass, 1996; Grace, 1998; Lomicka, 1998; Nagata, 1999; Laufer & Hill, 2000; and Yoshii, 2006) indicated that sentence-level definitions of words are helpful for learners.

What can be observed on the review on textual glosses is that there is no conclusive evidence to support if word or sentence type of glosses is more useful for gloss users. Clearly, more research is warranted for textual glosses, in particular word or sentence type of glosses.

In contrast to textual glosses, multimedia glosses incorporating text, sound, pictures and video have all been studied as glosses in different research settings. Researchers (Chun, Mayer & Leutner, 1998; Kost, Foss & Lenzini, 1999; Al-Seghayer, 2001; Yoshii & Flaitz, 2002; Mohamad Jafre, 2011 et al.) all investigated multimedia glosses and found them effective for vocabulary learning and comprehension.

As a result, once again, no definitive conclusion can be made on the use of the different type of glosses. To make matters denser, some researchers have also looked at various combinations of gloss which makes it even more difficult to compare the efficacy of the different modes of the gloss.

Generally, from the narrative on gloss use, it can be garnered that a combination of image-based and text glosses seem to be more effective than text only glosses for vocabulary acquisition (Al Seghayer, 2001; Plass, Chun & Mayer, 1998; Yeh & Wang, 2003; and Yoshii, 2006) The results from these studies fitted into Mayer's (2001) Multimedia Generative theory that postulates that information from both the visual and textual glosses would be processed in dual modes by the learners. Hence, there is more processing on the part of the learner and the integration of information provides for better vocabulary learning. There is also a contrasting view to this which will be explained next.

On the use of multimedia elements to gloss words, there is danger that the element of technocentricity may have a role to play in it. Researchers are awed with the technology available and set out to test them in various learning situations. The Cognitive theory of Multimedia Learning (Mayer, 2001) also posits that multimedia information may have deleterious effects on learning, more so when a single channel, for example, video is overloaded. This is evidenced by findings from some researchers who learnt that "apart from technocentricity, another consideration is that there is a certain threshold beyond which technology ceases to be effective" (Gettys, Imhof, Kautz, 2001, p. 92). The same researchers argued that there is a need to identify the learners' actual needs of the learners in using glosses.

Chun and Plass (1996) also provide several reasons for negative effects of multimedia gloss. They reason that readers may focus on the type of media used and not in the information in it, or its usefulness.

They recommend that multimedia materials be "designed as adaptive systems to support learners with different traits, such as learning preferences and cognitive styles" (p. 73). This should be done to provide learners with glosses in the modes which they need or prefer. Their conclusion is appealing and logical as the different types of media may not be beneficial to learners, unless due consideration is given to the learners' needs and preferences.

Researchers (Chun 2001; Davis & Lyman-Hages, 1997; Ercetin 2001; Laufer & Hill, 2000; Lomicka, 1998) showed that even when multimedia glosses are available, definitions or translations of words are the most commonly accessed annotations when the learners are given the freedom of choosing the glosses of their choice. This proved that definition type and translations are preferred mode of glosses. It also showed that the verbal medium is preferred to graphic, visual and audio media forms of gloss (Chun, 2001). Chun (2006) sums up this scenario aptly by writing that these studies have shown that definitions or translations of words are the most commonly accessed type of vocabulary gloss when learners are given the freedom to look-up glosses. She continues more pertinently by saying that verbal glosses do work well with other multimedia elements such as pictures or videos but "when given the choice, learners tend to prefer and use the simple translations of words" (p. 78).

This argument is carried on by Yanguas (2009) who carried a study on multimedia glosses and their effect on L2 comprehension and vocabulary learning. Although the study discovered that a combination of text and pictures had a positive effect on comprehension, the think-aloud protocols employed in the study revealed that some of the learners in the picture gloss group reported that the pictures "were not helpful and even distracting" (p. 61).

Apart from this, the reason for providing textual only gloss in this study is strengthened by Levy and Stockwell's (2006) observation that verbal or textual glosses, for example, were less complex to process as compared to images because less mental effort is expended to work out meaning.

From the review of the type of gloss, it can be seen that there is still debate on whether textual-only or multimedia is beneficial for learners. There is, however, a leaning towards the use of multimedia type of glosses. Nonetheless, there are instances where textual glosses may have the edge as the preferred mode of aid that learners need. Even between textual glosses, the literature has shown that the effectiveness of word and sentences glosses is both beneficial for learning vocabulary and comprehension. Hence, this study which utilizes textual-only glosses, breaks down these into word and sentence types with the aim that the findings can add more credible evidence to the use of such glosses in online reading contexts. The grounds for textual glosses are simple; given its simplicity and straight-forwardness, a lot of materials on the Internet can be glossed to provide better comprehension and opportunities for language development, in particular vocabulary development.

2.3.1.2 Language of gloss

Taylor (2006) in a meta-analysis study on the use of L1 glosses on L2 reading comprehension discovers that L1 glosses are effective means for comprehension. He argues for the case of L1 glosses as they act like a bilingual dictionary, provide fast access to learners, encourages look-up behaviour and more importantly, attentional resources of the learners can be used elsewhere, for example, reading for meaning.

It ought to be mentioned here that the studies reviewed (Aweiss, 1994; Bell & LeBlance, 2000; Hayden, 1997; Miyasako, 2002; Ko, 2005; Stoehr, 1999; & Taylor, 2006) were based on gloss use for reading comprehension rather than for vocabulary learning. Notwithstanding, these studies were reviewed as the glosses in those studies were vocabulary-based; what this means is that the glosses created provided meaning for difficult or unfamiliar words. The difference was the outcome of the studies which focused on reading comprehension instead of vocabulary development. Given the primal focus of vocabulary in reading, the review of the above-mentioned studies is not misplaced.

Another study on gloss language is work by Bell and LeBlance (2000) who suggest that L1 glossing would make learners more comfortable and it encourages positive anticipation of L2 reading, in the sense that the gloss provides vocabulary help in the learners' L1. Parallel to this, Aweiss (1994) found out from his sample of Arabic learners who had access to L1 glosses did better in L2 reading comprehension in the way of recalling more pausal units. Stoehr (1999) also discovered that participants in his study who had access to L1 glosses recalled a significantly higher amount of L2 text than those without glosses.

Hayden (1997) perhaps revealed an extreme reaction from learners when they hardly consulted other glossing options when there were L1 glosses available. Lomicka (1998) who studied university learners studying French found out that learners had a preference for L1 glosses in definitional-type of glosses. There were also researchers (Davis and Lyman-Hager, 1997; Lomicka, 1998; Nagata, 1999; & Li, 2010) who researched on using L1 in glosses. They have found out that using the learners' L1 in the gloss is beneficial for the learners. Cook (1999) also remarks that the use of L1 is gaining ground as a means of providing modified input.

He continues to state that the L1 use should be relooked as a value of translation in CALL where it can be used as a means of input modification.

In contrast, there were researchers such as Jacobs, et al. (1994) who found out that there were no differences between the vocabulary scores of the group who used English or Spanish glosses. Nonetheless, through questionnaire feedback, they discovered that half of the learners in their study preferred L2 glosses if the glosses were more comprehensible. This means that the glosses which are designed have to be at the right level for the learners to comprehend them in L2. Apart from that, researchers such as Laufer and Hill (2000) examined the use of both L1 and L2 in glosses and found no significant effects for vocabulary learning.

Miyasako (2002) carried out a study to find the effectiveness of L1 and L2 glosses in multiple-choice and single glosses. The study involved Japanese high school learners studying English. The researcher found out that L2 gloss groups outperformed the L1 gloss group on vocabulary learning. Ko (2005) who employed both qualitative and quantitative measures in her study discovered that the learners preferred to have glossed reading materials. She further found out that the L2 glosses showed significant effects on the learners' reading comprehension. The think-aloud protocol also indicated the L1 gloss helped facilitate comprehension.

Another researcher, Yoshii (2006) in his study also revealed that both L1 and L2 glosses were useful for incidental vocabulary learning in a multimedia context. In Yoshii's study, he had used 195 university students in a repeated measures design to investigate the effectiveness of L1and L2 glosses in four different gloss conditions: L1 text only, L2 text only, L1 text and picture, L2 text and picture. In sum, he researched the difference between L1 and L2 as well as text and picture glosses for vocabulary learning. Employing a set of vocabulary tests which he had designed as definition-supply and word receptive in a pre-test and post-test sequence, he found that both types

of glosses were useful for vocabulary learning although there was no significant difference of gloss type. Additionally, however, it appeared that pictures have an edge over text-only glosses in the definition supply test.

As for the language of gloss, it was discovered that L2 seemed to benefit the students more than L1 when a picture accompanied the gloss in the definition-supply test. Nevertheless, the L1 text only group had better retention rates compared to the other gloss conditions in the receptive task. Yoshii concluded that glosses are on the whole useful for vocabulary learning, regardless of type and language. He further remarked that the effect of L1 and L2 has to be examined more, perhaps including the learners' proficiency level as a research variable in future investigations.

In total, the results of the comparisons between L1 and L2 in vocabulary learning are still few and inconclusive. Ko (2012) surmises this scenario well by writing that studies on language of glosses are quite limited and important questions are still unanswered. She recommends that replications of studies are required to verify the findings of previous research and provide better insight into this area. Hence, with more studies, it can help define the role of L1 or L2 and its efficacy in gloss use and language learning. Therefore, there is a need for further investigation to find out the effectiveness of L1 and L2 glosses in vocabulary learning specifically and to examine which language is more beneficial to which learners and in what conditions.

Also, from the literature on language of glosses, it can be gathered that different type of languages of glosses comparing L1 and L2 have been investigated. The languages that have been studied so far are Spanish and English (Jacobs et al., 1994), Hebrew, Chinese and English (Laufer & Hill, 2000), Korean and English (Myong, 2005), Japanese and English (Miyasako, 2002; Yoshii, 2006), Therefore, this study hopes to contribute to the array of gloss studies, by adding another language which is *Bahasa Melayu* or Malay.

2.3.1.3 Proficiency level of gloss users

From the studies on gloss use which have been reviewed, it is revealed that some researchers have pointed out that the proficiency level of the learners does play a role in their look-up behaviour. For example, Ercetin (2001) states that, "second language learners interact with text differently based on their proficiency level and prior knowledge" (p. 70).

Jacobs et al. (1994) also found that the effectiveness of glosses varies according to the learners' language proficiency. They argue that glosses have a different impact on learners with different proficiency levels in the L2. In terms of vocabulary learning, they found out that higher proficiency learners recalled more if they had accessed to gloss words. On the other hand, Li (2010) who examined 20 Chinese ESL learners' vocabulary retention after they had read a text with and without access to computer-mediated English monolingual and English-Chinese bilingual dictionaries revealed that both computer–mediated L1 glosses and bilingual dictionaries were effective for learners with lower proficiency levels.

Miyasako's (2002) study disclosed that L2 glosses were more effective for learners with higher level of proficiency, while L1 glosses were effective or lower proficiency levels. On the same issue of language of glosses, Taylor (2010) in considering reading comprehension, observes that learners who prefer L1 glosses are from the lower proficiency group. Yoshii and Flaitz (2002) also considered the proficiency level of the learners who were at beginning and intermediate stages. They found out that there were no significant differences between the two levels in the rate of change between the immediate and delayed test scores. They conclude that the retention rate of vocabulary among the three groups did not differ significantly, although picture and text glosses may promote vocabulary learning in the short term.

With 106 Korean learners studying English as her sample, Ko (2005) carried out a study to find out how different types of glosses and the languages of the glosses (L1 and L2) affected the learners' reading comprehension. As for learners' proficiency, the findings pointed out that if the level of proficiency is high enough to understand the definitions in L2, then L2 glosses were more effective than L1 glosses.

Knight (1994); Yoshi (2006) and Abraham (2008) analyzed gloss use behaviour and have come to the conclusion that the language ability of the learners affects the way they use the gloss, that is, it has a role in the learners' look-up behaviour. Their interpretations of the gloss use have implicated the language ability of learners but these researchers have not delved deeper into the issue. Providing further evidence of the effect of proficiency and gloss use, Yun (2011) who carried out meta-analysis of 10 studies on gloss use on reading comprehension and vocabulary learning revealed that the variable of learner proficiency was found to be statistically significant as glosses made an impact to L2 vocabulary acquisition on beginning learners.

To make matters more complex in the use of glosses and proficiency, Abraham (2008) who carried out a meta-analysis of 11 studies on gloss use discovered that the language of the gloss, that is, L1 or L2 or a combination of both L1 and L2 played a significant role in learners' vocabulary learning for learners with different levels of L2 proficiency.

Added to the issue of language of glosses, Levy and Stockwell (2006) made known that recording learners' use of help facilities, for example, their access to online dictionaries or patterns of electronic aids is increasingly being related to the learners' language proficiency with the "goal of providing resources and guidance that meet the needs of the learners at a particular level" (p. 154). On top of this, Park (2002) also reported that there is evidence that different types of modifications may have different effects for learners at different proficiency level.

Thus, this researcher is confident that this research is significant in relating learners' look-up behaviour to their language ability.

Indeed, this study firmly believes that language ability of the learners plays a significant role in interactions or specifically in gloss use which could have an impact on second language acquisition, which in this case is the development of vocabulary knowledge. This point on learner variability is brought to a close with Ellis' (2012) comment:

> We have seen a number of studies that report a high level of variation in group gains, which points to the fact that learners differ in their capacity to learn through interaction. One possible explanation for this is that factors such as language aptitude, anxiety and motivation influence learners' ability and preparedness to "notice" form when engaged in meaning-focused communication. (p. 269).

2.3.1.4 Gloss use in the Malaysian context

A closer to home study on glosses was conducted by Liaw (2009). Using a sample of eight secondary school students in Melaka, Liaw set out to find out reading strategy awareness training of her respondents to empower online reading. Using Anderson's (2003) established Online Survey of Reading Strategies (OSORS) inventory, she found out that the respondents have used reference materials, that is, supporting tools such as the glossary and online dictionary to look up for meanings of words. Although the focus of her research is on the use of online reading strategies, it appears that online help in the form of glossary and online dictionary form part of the students' online reading strategies to empower their reading on the Internet which is another plus point for the use of gloss.

Another study of gloss in a Malaysian context was carried out by Faroogh Azari (2012). The objective of his study was to determine the effects of different textual glosses on reading comprehension in English as foreign language (EFL) learners in *Universiti Putra Malaysia* (UPM). From his sample of 76 post graduate learners who had read texts with different gloss aids in L1 (Persian), L2 (English), L1 and L2 glosses and no glosses, he discovered that learners using glosses outperformed those in the no-glosses texts in reading comprehension recall protocols. However, there were no significant differences between gloss groups. Furthermore, the learners preferred to read texts with both L1 and L2 glosses. This study lends weight to the benefit of glosses which can aid learners in reading comprehension. The other important point is that the language of glosses is immaterial as long as help is given in the form of glosses. The study is different from the current one is as it looks at reading comprehension, while the emphasis for the present study is on vocabulary learning. Nonetheless, it is noted that both the L1 and L2 used in the glosses were effective in enhancing reading comprehension. It will be interesting to see if the same benefits are gained for vocabulary learning in this present study.

Mohamad Jafre Zainal Abidin (2011) et al. carried out research comparing the use of multimedia annotations and printed textual glossary. The respondents in the research were 120 low proficiency Malaysian secondary school students. The researchers found out that students who had used the multimedia annotations had better recall and retention of new lexical items compared to the printed textual glossary group. This suggests that multimedia types of glosses are more useful than printed textual glossary for low proficiency students in learning vocabulary. Although the study had investigated multimedia and printed glosses, it is similar to this ongoing study as it had incorporated proficiency level as a factor in gloss use.

The other similar factors were its focus was on vocabulary learning and the design of the study which had post-test and delayed post-test corresponding to recall and retention of vocabulary items.

These studies were reviewed as they were research on the use of gloss in a Malaysian setting. The participants of the studies ranged from school to post-graduate students. The general finding from all the studies is that glosses are useful for students in their different presentations, for example, in printed or multimedia forms. It would certainly fill the gap in terms of the range and type of the participants in gloss studies in Malaysia, that is, from primary school children to postgraduate students. What is left to be seen is to find out if similar benefits can be obtained from the participants in the current study who are in their initial tertiary study.

2.4 Vocabulary knowledge

Ellis (2008) believes that SLA researchers do not deal with knowledge in describing second language acquisition. He is concerned that they skirt around this issue of knowledge by just reporting about "what learners have learnt or know without understanding the nature of knowledge they are investigating" (p. 427). With that in mind, this research briefly describes what knowledge is, in particular, implicit/explicit knowledge before delving further into what is meant by vocabulary knowledge.

At a basic level, 'knowledge' is defined in the Oxford Advanced Learners Dictionary as "the information, understanding and skills that you gain through education or experience" (2005, p. 821). Related to L2 knowledge are two types of knowledge which are implicit and explicit knowledge.

Gass (1997) explains that explicit knowledge is knowledge that learners are conscious of while performing a task, while implicit is the reverse, where there is no conscious knowledge of what the learners know while they perform a task. Ellis (2008) gives a comprehensive definition of what these two types of knowledge are.

According to him, implicit knowledge is intuitive, automatic and can be used in fluent unplanned language use, while explicit knowledge refers to knowledge which is conscious, declarative and accessible through controlled processing in planned language use. This issue of implicit/explicit knowledge is complex and contentious. The strongest contention is whether implicit/explicit knowledge forms a continuum or dichotomy of knowledge. The contention does not stop here but also discusses the notion of interface, that is, where explicit knowledge becomes implicit knowledge through practice.

It is also important to note that Ellis (2008) considers that L2 learners begin with explicit knowledge of the language. More importantly and relevant to language acquisition is explained by him as "explicit knowledge can contribute indirectly to the development of implicit knowledge by helping learners to notice linguistic forms in the input and comparing between what they have noticed and their own current interlanguage" (p. 423).

As can be observed from the preceding section, knowledge is a difficult construct to define. In vocabulary, the problem is more complicated as the definition of vocabulary knowledge varies among researchers. Clearly, vocabulary knowledge has been defined differently by different researchers. Nonetheless, it has to be defined and justly operationalised in vocabulary research in order for designs and outcomes on such studies are designed appropriately and the subsequent results can be interpreted meaningfully within the parameters of the definition.

Nation (2001) writes that vocabulary knowledge can be defined as the sum of interrelated subknowledges which involves knowing its "form, meaning and use" (p. 26). From these three spheres, knowledge of a word spawns knowing how the word is spoken, written, its form and meaning, concept and referents, its associations, grammatical functions, collocations as well as its constraints on use (please see Appendix B for "What is involved in knowing a word"). Furthermore, vocabulary knowledge is looked upon as a continuum comprising several layers.

The first layer can be considered as superficial familiarity of the words by researchers such as Faerch, Haastrup, and Phillipson (1984) and Palmberg (1987). At the next level is one of the most common distinctions of vocabulary knowledge which is "receptive and productive knowledge" (Schmitt, 2010, p. 80).

Nation (2001) elaborates that receptive knowledge is also termed as passive knowledge and this kind of knowledge is needed to deal with words in listening and reading. On the other hand, productive knowledge is called active knowledge and it is needed to use word in speaking and writing.

2.4.1 Vocabulary development

It has been explained in the previous section that vocabulary knowledge is complex. This complexity is increased when vocabulary learning is discussed. Vocabulary researchers such as Schmitt (2010) and Henriksen (1999) state that vocabulary learning is incremental in terms of acquiring an adequate vocabulary size as well as acquiring individual lexical items. Schmitt continues to state that the vocabulary learning is "incremental in a variety of ways because some types of word knowledge are established before others" (p. 20).

For instance, he elaborates that a learner would probably have an initial grasp of a word's basic meaning before moving on to acquiring other types of knowledge such as word associations or collocations.

At another related level, Henriksen (1999) proposes that learners develop vocabulary knowledge on a continuum as against a dichotomous of "know" versus "do not know" a word. According to Schmitt (2010), a well-researched vocabulary knowledge area is in meaning. In this aspect meaning may include receptive and productive vocabulary knowledge. Although this distinction of receptive and productive vocabulary knowledge seems straightforward, there is debate as to whether this knowledge constitutes a continuum or is dichotomous.

Some researchers like Melka (1982) states that it is a continuum, while others such as Meara (1997) considers it as different and that there is no natural progression from receptive to productive knowledge. Regardless of their stand on this issue of whether receptive and productive knowledge as being on a scale or at separate poles, they all accept that such a dimension exits (Pignot-Shahov, 2012). Furthering this line of inquiry, researchers such as Laufer and Paribakht (1998) and Laufer and Goldstein, (2004) report that receptive mastery generally develops before productive mastery. Therefore, what can be gathered here is that vocabulary learning is a gradual process and that different types of vocabulary knowledge are developed at various points in this process of developing vocabulary knowledge.

2.4.2. Learning vocabulary

Schmitt (2002) writes that for most first language vocabulary learning, a large part of the input comes from listening and reading. The same can also be applied to second language learners, where Read (2004) states that second language learners

acknowledge the importance of vocabulary in the target language. Furthermore, second language learners need to be exposed to different modalities such as written or listening to get input in order to acquire new vocabulary. Two different perspectives are put forth next to discuss vocabulary learning.

2.4.2.1 Nation's Vocabulary Learning Theory

From Nation's (2001) broad research on vocabulary learning, he suggests that there are three stages which are required for vocabulary acquisition. They are noticing, retrieval and generative. The first is Noticing where the learner notices that there is a word which he or she is unfamiliar with. The second is Retrieval where there is a possibility of memorizing of words so that they can be retrieved by the learner, and the third is generating the vocabulary retrieved. Nation makes a distinction between retrieved vocabulary knowledge: receptive retrieval and productive retrieval. Receptive retrieval is when learners perceive the form and retrieve the meaning, while the opposite is true for productive, that is, learners have the meaning and retrieve the form. Nation puts across that both types of retrieval are important, however he feels that productive retrieval is better for vocabulary learning. The final concept is Generative where learners are able to use the words learnt in a different grammatical form, in different contexts or with a new meaning (Alum, 2004).

2.4.2.2 Laufer and Hultsjin's Involvement Load Hypothesis

Another vocabulary learning model which can also be applied to the interaction between the computer and the learner is provided by Laufer and Hultsijn (2001).

They have termed it as the Involvement Load Hypothesis which posits that tasks which induce higher involvement from the learner. It works on the premise of the learners' level of processing, that is, when the learner processes the vocabulary more, the eventuality for the vocabulary to be learnt and retained is increased. Laufer and Hulstijn (2001) suggest that the involvement of vocabulary learning has three components. They explain that the three components are: Need, Search and Evaluation. Need is the requirement to understand a linguistic feature in order to perform a task, for example, need for a meaning of a word in a reading comprehension task. Search is to look for the meaning, which could be looking up the word in a dictionary or even use a gloss. Finally, Evaluation is to evaluate if the word can be suitably used in certain contexts. This Evaluation stage is more a cognitive type of interaction, one which Chapelle (2003) terms as "Intra" that is within the learner's mind.

The involvement load can be triggered by the learner first noticing the lexical item, the amount of time spent engaging with the lexical item in terms of interaction thus increasing the engagement with the lexical item. These are among the factors which Schmitt (2010) observes that can facilitate vocabulary learning. He adds that "the more a learner engages with a new word, the more likely he/she is to learn it" (p. 26).

The preceding section has discussed two perspectives on vocabulary acquisition. The discussion shows that there is increasing understanding of vocabulary acquisition system. However, it should be kept in mind that it is difficult to define any theory to vocabulary acquisition. This is articulated by Schmitt (2010) as "overall acquisition system is far too complex and variable for us to comprehend in its entirety, and so it still eludes description" (p. 36). Despite this, he continues that the way forward in overcoming this gap in the area is to have large numbers of studies with different methodologies before some key developmental patterns in the area can be identified.

2.4.3 Vocabulary attrition

The earlier sections in the chapter have shown that vocabulary learning is certainly a dynamic process. It is seen that vocabulary development is possible however as in most kinds of learning, it is common that there will be instances of attrition (Schmitt, 2010). In fact, according to him vocabulary knowledge is more susceptible to be lost when compared to other linguistic elements such as phonology or grammar. He reasons that this is because vocabulary is made up of individual units instead of a series of rules as in grammar, hence the tendency for backsliding to occur is greater.

In relation to the type of vocabulary knowledge which is prone to attrition, there has been mixed results. Researchers such as Cohen, Olshtain, Weltens & Grendel, (as cited in Schmitt, 2010) have reported that receptive vocabulary knowledge loss is higher than productive knowledge. On the other hand, Schmitt (1998) in his own longitudinal study on tracking vocabulary acquisition found the pattern reversed. It can be interpreted here that vocabulary knowledge gain or attrition is a phenomenon which is largely fluid because of its many intercepting variables.

2.4.3.1 Memory

When attrition is discussed, it would not be complete if the aspect of memory is not touched on. Psychologists describe three kinds of memory depending on the time something is initially presented and moving on to the retrieval stage. The first stage is where the initial encounter with something to be remembered is termed as the sensory store. This is where the receptor organ identifies the item encountered.

The second stage is called the short-term memory where attention is maintained after the initial item identification phase; it is whatever that is holding one's attention. As the name suggests, in the short-term memory, nothing is stored for long here. Nonetheless, retention and retrieval in short-term memory is immediately accessible.

Another term which is used to describe short-term memory is working memory. Gass and Selinker (2008) define working memory as "referring to the structures and processes that human use to store and manipulate information" (p. 250). They make the distinction that the difference between short-term memory and working memory is that in the latter the focus is on the manipulation of information while short-term memory is solely for storage. Apart from these, there is long-term memory where there exists high capacity for storage. Its accessibility or retrieval depends on how the knowledge is organized in the long term memory.

It is apparent that there is a relationship between learning a language and memory. Gass and Selinker (2008) explain that the ability to remember information about a language may bring about language learning. In other words, the higher a language learner's capacity of remembering, the more success he/she will have in language learning. Another aspect which is related to working memory is the language learner's proficiency level. It has been shown that there is a correlation between working memory and proficiency. There is significantly weaker correlation with lower proficiency learners than with more advanced learners (Gass & Selinker, 2008).

2.4.4 Measuring vocabulary knowledge

Clearly, it can be seen that research on vocabulary is a complex matter, whether it is in defining the vocabulary construct and even pinning a theory to learning vocabulary.

This complexity continues when it comes to measuring vocabulary knowledge. It is no easy task to measure vocabulary (Nation, 2001). This claim by Nation is aptly described by Schmitt (2010) who states:

> It is virtually impossible to measure all the word-knowledge aspects for words for a least three reasons. The first is that many of the word knowledge aspects do not have accepted methods of measurement. A second reason has to do with time…a test battery for word would be extremely unwieldy and time consuming. A third reason is related to the difficulty of controlling for cross-test effects. (p. 79).

Nevertheless, efforts have to be made to provide justifications for the types of tests used and to explicitly state what the tests measure and to what extent. A good starting point would be to define the construct vocabulary as explained in the earlier section. The operationalising of the vocabulary construct is to make a distinction between the types of vocabulary knowledge tested. As there can be different types of vocabulary knowledge there is a need for multiple tests to measure the varied aspects of vocabulary knowledge. Another use of multiple tests for vocabulary is because a single test would not be able to measure every aspect of word knowledge (Milton, 2009). Different types of tests are developed to capture the different types of vocabulary knowledge of learners with different degrees of sensitivity. These tests can be developed along the lines of discrete-selective-context-independent dimension developed by Read, 2000 (p. 9).

This is shown in Figure 2.3.

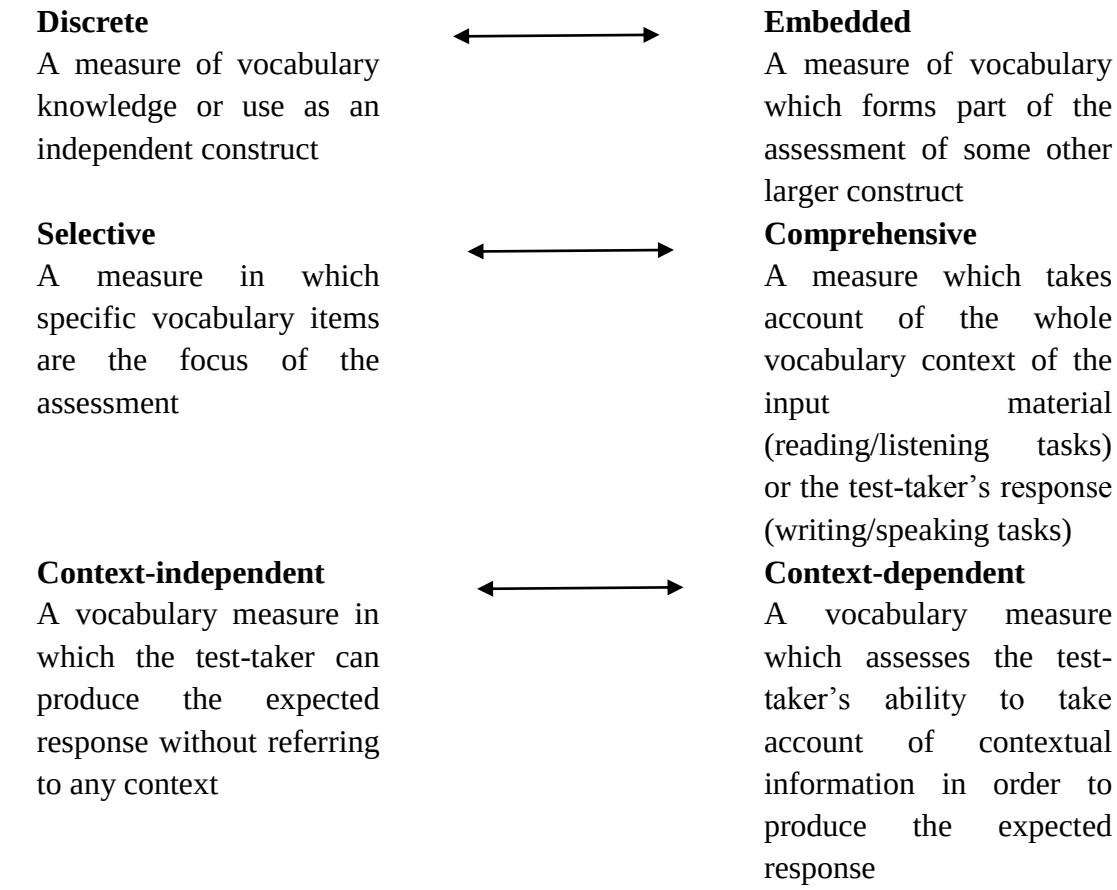


Figure 2.3: Dimensions of vocabulary assessment.

2.4.5 Vocabulary and reading

The importance of vocabulary to reading achievement, more specifically to reading comprehension has been established. Knowledge of word meanings and the ability to access that knowledge efficiently are recognized as important factors in reading and listening comprehension, especially as learners progress to higher or tertiary level (Chall, 1983). Schmitt (2000) remarks that the bulk of words that learners need to know will have to be acquired not by formal vocabulary study but through non-teaching activities, that is exposure to language, where reading is an important source.

This same sentiment is echoed by Grabe (2004). He observes that in an L1 environment, vocabulary is acquired mainly through spoken input, however, for the L2 learners vocabulary acquisition often occurs more through written texts. This observation strengthens this research project in emphasizing vocabulary knowledge acquisition through online reading via the Internet which has become an obvious source of English for today's learners. This is apparent even to a casual observer without referring to any empirical evidence. As for the importance of vocabulary in reading, Laufer (1997) claimed that "no text comprehension is possible, either in one's L1 or L2, without understanding the text's vocabulary" (p. 20).

More evidence to show the critical role of vocabulary in reading is provided by Hu and Nation (2000) who report that knowledge of 80% of the words is the minimum threshold to comprehend a text, but most learners need to know around 98% of words in order to read independently. Therefore, vocabulary load is seen as a significant predictor of text difficulty. Haynes and Baker (1993) make the conclusion that the most significant handicap for L2 readers is not lack of reading strategies but insufficient vocabulary in English. This indicates that the threshold for reading comprehension is to a large extent, lexical.

Read (2004) also comments that second language learners are aware that limitations in their vocabulary hinders them in their communication in the target language as "lexical items carry the basic information load of the meanings they wish to comprehend and express" (p. 146). This is the reason why Read justified that vocabulary is important for learners to acquire as compared to other features of the language. Furthermore, Grabe (1991) notes that "virtually all second language reading researchers agree that vocabulary development is a critical component of reading comprehension" (p. 392).

Also in this track, Cook (2001) claims that "learners are often acutely aware of their ignorance of vocabulary as compared to their unawareness of their ignorance of grammar and phonology" (p. 66). It is clear then that vocabulary is one of the main factors in reading and learners are aware that the lack of vocabulary is one of the main reasons for their lack of comprehension. As Ellis (1999) points out, "often the negotiation of meaning focuses quite explicitly on lexis.

The question then is how negotiation of or interaction with meaning can help learners acquire vocabulary? Ellis (1999) suggests that it can do so by (1) "increasing the overall quantity of the input and (2) through elaboration that provides clearer meaning for the learners" (p. 51). All this evidence for the importance of vocabulary indicates that there is justification for the current study to investigate vocabulary knowledge development in the light of gloss use within an online reading task.

2.4.5.1 Online Reading

Research in second language (L2) reading dates back to the 1960s. Traditionally, it involves factors in the bottom-up approaches that are letter-receptive, word-receptive, and the top-down approaches involving reader's background knowledge and reader's use of strategies. The combination of the two approaches in L2 reading is called the "interactive approach" (Bernhardt, 1991).

The Internet provides new text formats, new purposes for reading and new ways to interact with information (Coiro, 2003). Online reading contexts have an intriguing irony on the aspect of reading as being interactive which was brought up by Reinking and Leu (1996). They point out that the metaphor lies in that in order for interaction to be present, there has to be active process from the two parties involved in the reading process. The reader and the text have to interact with one another.

They present their case that this interaction does not occur while reading traditional texts. They suggest that reading in an electronic environment brings out the real meaning of "interactive". According to them, an electronic reading environment will provide the reader with the opportunity of interacting with the text, for example, with the use of electronic gloss. Therefore, the interaction that most reading theories suggest, actually takes place not only within the reader, internally, but also externally between the reader and the text. The text is able to respond with the requests made by the reader. These requests could be in the form of definition of unfamiliar words, grammatical aid, provide background knowledge or even provide an explanation for a difficult concept. Online readers can leave the online text to access these aids and resume reading. In total, it can be said that online reading is a non-linear and discontinuous process (Youngs, Ducate & Arnorld, 2011). Therefore, this insight is useful for the current study, as it can be seen that interaction is central to the reading process.

2.6 Conclusion

This chapter underlines the theoretical framework of the study, focusing on the notions of input and interaction in SLA. Besides this, other related aspects of SLA such as noticing, negotiation and output are also spelt out in relation to how they can be applied to a CALL setting; focusing on the use of computer-aided glosses The intricacies of what constitutes vocabulary knowledge, its development and measurement are also laid out.

It is also clear from the literature review that there are two directions in the study. The first is from the standpoint of the input-interaction view where the roles of input in the form of an online text and the glosses which provide the modified input of the study. Together, they form the basis for learner-computer interactions in a CALL context.

Next, the research also looked at the suitable conditions for vocabulary development brought about by these learner-textual gloss interactions in an online reading environment. The literature review then draws up these elements together to provide direction to the study by examining the type and language of computer gloss interactions which can promote vocabulary development. By using a quasi-experimental design, this study compares how learners with different language proficiency levels interact with the glosses in the context of vocabulary development. The design, set-up and procedure of the experiment are further discussed in the next chapter of the thesis.

CHAPTER 3

METHODOLOGY

3.1 Introduction

This chapter outlines the research methodology used in the study. It begins with a sketch of the design followed by descriptions of the sample population, research instruments and procedures that were used to collect and analyse the data.

Being mindful of the call by Chapelle's (2001) for a more robust methodology which can clearly show how the learning of language can occur in a CALL domain, this study used a quasi experimental design to provide empirical evidence of how modified interactions between learners and a computer-aided vocabulary gloss were able to help learners expand their vocabulary knowledge. In order to find out if learners had developed their vocabulary knowledge through the interactions, the study used a quasi experimental, group, pre-test, post-test design. As in any experimental design, this study has some form of comparison, that is, comparison which is provided by the different language abilities of the learners in the experimental group. The within group comparison is provided by the different tests in the experiment which was carried out at pre-, immediate and delayed stages of the study.

The study's primary focus was to examine closely the interaction of the learners with a computer-aided textual vocabulary gloss. The interactions by the learners with the gloss would form the basis for obtaining modified input in the form of the glosses which is a feature that is claimed to enhance second language acquisition which may be triggered by lexical problems (Ellis, 1999).

3.2 Purpose of study

The basis of the inquiry of the study was to investigate how learner-computer textual gloss interactions led to vocabulary knowledge development. From here, this basic question splintered into two threads. The first was to investigate the type of learner-computer interactions through the look-up behaviour of learners by obtaining modified input in the form of word or sentence glosses and its language, whether in L1 or L2. The next thread was to explore the extent to which language ability of the learners was a factor in their look-up behavior that allowed them to expand their vocabulary knowledge. Essentially, the study examined if modified input in the form of glosses did facilitate vocabulary expansion and what were the conditions for this development, and whether there was vocabulary knowledge gained and sustained.

3.3 Research questions

From the purpose above, it leads to the following research questions:

1. What is the clicking behaviour of the learners in the different gloss conditions?
2. (a) Which specific type of modified input, that is, word or sentence and in learners' L1 or L2 in the interactions facilitates (i) perceived vocabulary knowledge (ii) productive vocabulary knowledge and (iii) receptive vocabulary knowledge of learners with different language proficiency levels?

   (b) Is the knowledge maintained over time?

3. Is language proficiency of the learners a factor in determining how the learners interact with the glosses and subsequent effect on vocabulary knowledge?

3. 4 Research design

The research design was a quasi-experimental, pre-test-post-test design. This is in line with Ellis (2008) who emphasizes that "there is move towards studying input/interaction in laboratory settings, experimental and quasi-experimental methods have been used to investigate the effect of specific variables on input and interaction" (p. 207).

In this research design the learners were stratified according to their language proficiency based on their *Sijil Pelajaran Malaysia* (SPM) English grade (please see Table 3.1). Once they have been stratified into high, mid and low proficiency levels the learners were then randomly assigned into the four gloss conditions: word *Bahasa Melayu* (WBM), sentence *Bahasa Melayu* (SBM), word English (WEN) and sentence English (SE) thus forming the four experimental groups where learners were given the same online reading input.

To measure how the interactions affect vocabulary acquisition, vocabulary pre-test and post-tests were designed and administered to the learners at set intervals, namely, pre-experiment, immediately after experiment, and post experiment. Gass and Mackey (2007) assert that such designs are able to investigate the link between interaction and learning which would then consolidate the value of the interaction data.

The data collected was analysed using both descriptive and inferential statistics. The mean and standard deviation of the vocabulary test scores were analysed and presented graphically. Next, the data was analysed using Mixed ANOVA in examining the within-subject effect of time and between-subject effect of type and language of gloss. The next section discusses the variables in the study.

3.5 Variables

Before the variables in the study are described, the unit of analysis for this research would be the interactions between the learner and the computer. The primary independent variable in this study was the gloss that provided the modified input for the students. In this modified input, four different types of glosses were provided. They were word definitions in L1, sentence level meanings in L1, word definitions in L2 and sentence level meanings in L2. The use of the gloss was measured by the number of clicks the learners made on the targeted words which then revealed the type and language of the gloss accessed. In short, the type of gloss, language of gloss, and the number of times a learner clicked on the gloss were documented and tracked by a tracking device incorporated in the online reading text. The online reading text and its tracking device are described under the sub-section 3.8.1.5 of this chapter.

The dependent variables were the vocabulary test scores of the learners. In the study, these consisted of a pre-test, an immediate post-test, followed by a delayed post-test. The tests were constructed to capture the pre- and post- vocabulary knowledge of the learners. Initially, a pre-vocabulary knowledge test was carried out. Once this was done, two other types of tests were designed to measure receptive knowledge and productive use of the target words. Issues related to the development of the tests will be discussed later in this chapter under the sub-section 3.8.2.

The focus in this research was on the aspect of form and meaning. For this research the basic definition of knowing a word involves "form, meaning and use" (Nation, 2001, p. 26). Furthermore, vocabulary knowledge was looked upon as a continuum comprising several layers. The first layer can be considered as superficial familiarity of the words by researchers such as Faerch, Haastrup, and Phillipson (1984) and Palmberg (1987). For this study, this first layer was termed as perceived vocabulary knowledge as it stemmed from the students' own perception of their knowledge on the target words. At this first level, the learners self-rated their own knowledge of the target words.

At the next level, vocabulary knowledge was classified as receptive and productive type of knowledge. According to Schmitt (2010) one of the most common distinctions of vocabulary knowledge is receptive and productive knowledge. Pointedly, the receptive vocabulary knowledge was concerned with what meaning does the word form signal, while from the productive vocabulary knowledge angle, looked at what word form can be used to express this meaning. This led to the Operationalization of vocabulary development as to "know" a word and its meaning at three levels: perceived, receptive and production knowledge. The knowledge is evidenced from the results of the different vocabulary test scores that were administered at three points in the experiment.

According to Mackey and Gass (2005), an intervening variable is a variable which was not considered in a study on two grounds. Firstly, a researcher may not have considered the possibility of its effects and secondly it cannot be specifically identified. In the case of this research, the intervening variables are consciousness and attention of the learners towards the target words which have been made salient. Although it has been stated that the clicking of the target word may indicate the element of consciousness and attention, the study is unable to record these elements.

These elements of consciousness and attention cannot be easily understood as they are more of a cognitive-type of acts of the learners. In other words, clicking on the target words may not necessarily signal consciousness and attention. As a result of this, there could be implications to the results and the researcher was careful with the interpretation of the results of the study in particular the clicking behaviour of the learners in the study.

The other intervening variable of the study is the developmental level of the learners. This particular aspect is important in interaction research as researchers agree that the learner has to be at the right developmental level to be sensitized to benefit from the input, noticing and subsequent interactions to aid in language acquisition.

Mackey and Gass (2005) further define a moderating variable as one which may modify the relationship between the independent and dependent variables. One of the moderating variables which may have affected the way the learners use of the gloss was the learners' L2 proficiency. In this study, the learners' L2 proficiency was the moderating variable as it may have an effect on the learners' look-up behaviour.

To control the effect of this variable, the researcher categorized the learners' language proficiency by segmenting learners' proficiencies into three categories of learners' L2 ability. These categories are "low", "medium" and "high". This is based on the learners' *Sijil Pelajaran Malaysia* (SPM) English grade. The SPM is a national examination which is taken by all students in the Malaysian education system after approximately 11 years of school. Students take the SPM in form five and English is one of the subjects which is offered to the students. The grades in the subject would be an indicator of the learners' language proficiency. The configuration of the SPM grades into "low", "medium" and "high" for use in this research is shown in the Table 3.1.

Table 3.1: SPM Grades corresponding to learners' level of proficiencies

Grades	Category Values	The learners proficiency levels accorded in the study
A+	Highest Excellent	High Proficiency
A	High Excellent	High Proficiency
A -	Excellent	High Proficiency
B+	Highest Distinction	Mid-Proficiency
B	Higher Distinction	Mid-Proficiency
C+	High Distinction	Mid-Proficiency
C	Distinction	Mid-Proficiency
D	High Pass	Low Proficiency
E	Pass	Low Proficiency
G	Fail	

The other variables in this study were the vocabulary knowledge of the learners before the experiment, L1 background and number of years exposed to the L2. (Please see Appendix D for a diagram depicting the variables in the study). These variables were controlled by having a pre-test to measure the learners' vocabulary knowledge before the experiment. As for their L1, only learners whose L1 is *Bahasa Melayu* were included in the sample of learners for the experiment. As an added control, only learners in part one were selected to be in the experiment, hence, their years of exposure to L2 in schools would be about 11 years for all of them. Table 3.2 presents how the learner variables in the study were controlled, while Table 3.3 presents how the text or task variables are controlled.

Table 3.2: Learner variables that were controlled

Variable	How the variable was controlled
L1 proficiency	SPM English grades were categorized as High, Mid and Low
Vocabulary knowledge of target words	A vocabulary knowledge test was designed
Age	All learners were in part one, hence their ages were in the range of 19-20 years old

Table 3.3: Task variables that were controlled

Variable	How the variable was controlled
Text type, length	Narrative 483 words
Target words	Selected verbs which were unfamiliar to learners
Target words highlighted	A different colour was used for the target words
Readability of text	Lexical density Density of unknown words

3.6 The Participants

The students for the research were in part 1, the first semester, of their diploma studies in *Universiti Teknologi Mara* (UiTM). They underwent UiTM's first semester English course after studying English in both primary and secondary schools for a total of 11 years. They had been accepted to study in UiTM after completing their *Sijil Pelajaran Malaysia* (SPM) examination. The SPM is held after 11 years of primary and secondary education. For most of the learners in UiTM, they come from the national school education system where the primary medium of instruction was in *Bahasa Melayu* and English was taught as a subject. In order to gain admission into UiTM, a credit in English in the SPM was a requisite for the learners.

Whilst most institutions of higher learning in Malaysia retain *Bahasa Melayu* as the medium of instruction in most of their courses, the situation in UiTM differs as all its courses are conducted in English. To ensure that UiTM learners are able to cope with their studies in English, the university runs mandatory English courses for its learners. All learners have to pass these courses in order to graduate.

The participants for the research were from the Faculty of Survey, Planning and Architecture, UiTM Perak. This faculty is the largest faculty in UiTM Perak in terms of the number of learners.

Therefore, by drawing on a sample of learners from this faculty may be more representative of the population in UiTM Perak. Participants in this research were in their first semester of their Diploma in Quantity Surveying programme. They were taking a compulsory course in English, BEL120 which is a foundation course in English, focusing on grammar, reading comprehension and writing. BEL 120 forms the first English course in a three-semester English programme in UiTM. They were all Malay learners and have *Bahasa Melayu* as their L1. There were 117 learners at the start of the study and their ages were between 18 to 19 years old. Ninety-nine learners finally took part in the study as some of the students did not complete the whole set of vocabulary tests and some students' interaction behaviour was not documented.

3.7 The pilot test

A pilot test was conducted to check if the website and its gloss were working as expected. Thirty students took part in the pilot test. Similar to the participants in the the research, these students were also in their first semester of studies at Universiti Teknologi MARA (UiTM). They too were of mixed ability in their English proficiency based on their Sijil Pelajaran Malaysia (SPM) English results. These selection methods were used to ensure that the students in the pilot test closely resembled the participants of the study. The students were divided into two groups, one control and the experimental. The experimental group had access to the online text with the glosses, while the control group did not. As expected, the results from the tests showed that the experimental group performed very much better in the vocabulary tests.

This indicator plus the literature on the positive use of glosses for comprehension and learning (Nation, 2001; Myong, 2005; Watanabe, 1997), the researcher decided not to include a control group in the study.

This lack of the control group would not make the experiment less robust as it is already clear that there would be benefits in using the gloss. The focus is solely on the groups which had access to the gloss and what kind of glosses would benefit the students in their vocabulary knowledge.

In the pilot test, when the students clicked on the target words, they were presented with the whole range of glosses, that is, at word BM, sentence BM, word EN, and sentence EN. In short, the whole range of possible types of modified input was made accessible to them. The students had to choose their preferred type of gloss. This meant that it was up to the students to choose the type of modified input offered to them.

The students also had to undertake the different types of tests; namely a perceived vocabulary knowledge test, word receptive and production tests at three points: pre-test, immediate and delayed. The format of the pre-test was in the form of a definition-supply test. In the test, they were required to provide the definitions of the target words in either the L1 or L2. Next, the word receptive test required the students to choose the meanings of the target words, while for the word production test, students filled in gap with the correct words.

The online text and its glossed words worked as expected. However, the results from the look-up behaviour posed a problem. The variance between the types of looked-up glosses was too wide and no pattern of usage was seen. A decision was made to change the design of the research in terms of access to the glosses. Instead of allowing the students to choose their preferred glosses, that were WBM, SBM, WEN, SEN, students now had access to specific glosses. In other words, the interaction of the students was limited to one type of gloss. In this way, the effectiveness of each type of gloss could be measured directly.

The other significant impact from the pilot test was the definition-supply test. In this test, students supplied the meaning of the target words in either L1 or L2. The tests were scored by two lecturers. The inter-rater reliability between the two examiners was low. This partly could be attributed to the nature of the answers which could be in either the L1 or L2 which made interpretations of the answers subjective. The other drawback was that the test was time consuming. The word receptive and production test did not pose a problem. However, the test effects were significantly felt if the tests were given as pre-tests. The students scored very high marks for the immediate and delayed test which could be the result of the pre-test where the students were already made aware of the nature and purpose of the tests. Thus, two decisions were made on the type of tests to be used in the study. Firstly, the definition-supply test will be replaced with a self-rate vocabulary test. Secondly, the word receptive and production test will only be administered as immediate and delayed tests to reduce the test effects.

3.8 Data collection tools

In this section the data collection tools is discussed. It has to be made clear that in this study the instruments used for data collection can be classified in two ways. One was the means to collect interaction data and secondly the instruments that provided the evidence for the effect of gloss use on vocabulary knowledge development.

The means were the online reading text with its highlighted and glossed target words, and the tracking device. While the vocabulary tests were the instruments which gathered evidence for vocabulary development.

- Means: - The text
  - Target words
  - The gloss
  - Tracking device
- Instruments: - The vocabulary tests

Data for the study were also collected at two levels. The first was data which were related to the interaction of the learners with the gloss.

The online reading text with the highlighted and glossed words was presented to the learners online at www.mohdalionline.com/gloss. With a built-in tracking device in the online reading text, look-up behaviour of the learners was documented. It contained information such as the type of gloss clicked and the language of the gloss.

At another level, data were in the form of vocabulary test scores. The purpose of the tests was to show evidence of vocabulary development from the interaction of the learners with the gloss. Three different types of tests were constructed for the study. In section 3.8.2 "Data collection instruments", the instruments are described according to their roles in the data collection process. Table 3.4 summarizes the means, instruments and type of data found in the study.

Table 3.4: Means, Instruments and type of data

Data collection tools	**Type**	**Data**
Means	Online text	Target words clicked
	Gloss	Type of gloss looked-up: -word -language
Instruments	Pre-test Post-tests	Vocabulary scores

3.8.1 The text: "A Scary Night"

The source of the text, its purpose and details of the text in the form of its content, readability, and how the target words were selected for glossing are explained in this section.

In selecting the appropriate text, several important criteria had to be borne in mind. Firstly, the text used had to have a similar conceptualization of the phenomenon being studied which in this case, would be the gloss look-up behaviour of learners and vocabulary development.

Secondly, the text also had to be consistent with the current research and it had to be at a suitable level for university ESL learners. The selected text formed the input for the learners. It was a narrative text titled "A Scary Night" (Please see Appendix C). The text was taken from Professor Dr. Makoto Yoshii from the Perfectural University of Kumamoto in Japan. Permission to use the text as well as the tests for Yoshii's (2006) study was obtained through personal e-mail communication (1 November 2010). Also, the study was interested to use a text which had been used in a related and similar study to enable findings to be compared. The researcher is aware that direct comparison of findings may be limited as the gloss configurations of the two studies (Yoshii's, 2006 and the present study) are different. However, some form of comparison can be made as both studies utilised textual form of glosses.

A narrative text was used instead of an expository one because Abraham (2008) reveals that computer-mediated glosses had a medium effect on comprehension and vocabulary learning of (m=.64) and a large effect for narrative texts (m=.91). This means that a narrative text offers more opportunities for comprehension as well as vocabulary learning.

Yoshii (2006) had used the text for a similar experiment where he had glossed words in the reading text as target words which were considered difficult for his ESL learners. The words were all verb forms and were glossed in three different conditions. The gloss used in Professor Yoshii's study consisted of three different types. They are text-only meanings, picture-only and text and picture.

In this current study, the same text was used in this study but the difference was the way the target words were glossed; only textual gloss was provided in two languages that were the learners' L1 and L2. The rationale for providing textual gloss and the use of L1 and L2 in glosses was provided in Chapter Two. (Please see sections 2.3.1.1 and 2.3.1.2).

Before the details of the text are discussed, the gist of "A Scary Night" is given. The text was a story about a student who, while studying for his examination, witnessed his neighbour's house being burglarized late one night. The story continued with him reporting the matter to the police and with the burglars being caught subsequently. The story ended with a coincidental twist.

3.8.1.1 Suitability of the text

A brief description of the surface structure of the text is now given before moving into more details of its readability in terms of lexical density and density of unknown words. The surface structure is an important consideration as it forms the input for the learners. Ellis (1999) claims that "learners need input that is adjusted to their level (p. 249). As input, the text has to be at a suitable level for the learners to comprehend it with the exception of the targeted lexical words. The elements of the surface structure are given in Table 3.5.

Table 3.5: The breakdown of the structure of the online reading text

Item	Number
Number of characters (without spaces)	1,874
Number of words	480
Number of sentences	72
Average number of characters per word	3.90
Average number of syllables per word	1.29
Average number of words per sentence	6.67

The length of the text is also a factor in the surface structure; a short text would be inadequate as it would not provide enough information for comprehension. On the other hand, a lengthy text would take considerable time for the learners to read and could affect the overall experiment. The text written by Professor Yoshii for his experiment was 480 words long which fitted into the current study's learners' English course where they were exposed to reading texts of similar length.

3.8.1.2 The Readability of the text

Ellis (1990) states that input complexity can have an impact on both comprehension and acquisition of new words. It is then important that the readability of the text be scrutinized in order for the text to be pitched at the right level for the learners. The study employed two ways to establish this. One was by expert agreement and the other was by lexical density, and the density of unknown words which are explained next.

Two experienced lecturers who taught the first semester English course In UiTM Perak were consulted on the suitability of the text for the semester one learners. Both of them agreed that the text was at a suitable level for the learners.

Once this was established, the researcher subjected the text to a procedure where its lexical density was measured.

(a) Lexical density

Lexical density shows how much information is carried in a text. Lexical Density is measured by breaking down the text into content or lexical words and function words. The lexical density of a text is measured by dividing the number of lexical words with the other words in the text and multiplying it by 100. In the text used for this research, the formula is worked as:

225 (lexical words) ÷ 480 (total number of words in text) X 100 = 46.87%

A high percentage indicates a large number of information-carrying words and a low lexical density reflects the opposite. The lexical density of the text used is 46.87% which means that the text has a somewhat fair distribution of lexical and function words which can be interpreted as having mid-density which would not pose comprehension problem for learners as would a lexically dense text. Hence, the text used is appropriate for the intended learners.

(b) Density of unknown words

The density of unknown words is measured by establishing the ratio of new words to total words in a text (Ellis, 1999). Nation (1990) refers to a study done by Holley (1973) who tried to measure the optimal level of new-word density for written text.

It showed that vocabulary learning can occur even though new vocabulary density of one new word is to fifteen (1:15) known words. In the text used in the study the ratio is 1:36 indicating that the text is not dense as compared to the one cited by Holley. The ration recommended by West (as cited in Ellis, 1999) is 1:50 which he stated is appropriate for supplementary readers. In the text used, the density of unknown words was between the ranges of 1:15 and 1:50 which make the text readable for the participants of the study.

The other more direct measure of indicating the suitability of the text was to determine syntactic complexity of the text for the target learners (Al-Seghayer, 2005). In order to determine this, the type of sentences that made up the text was analyzed. There were 72 sentences in the text. The mean length of the sentences was approximately seven words. What this translates into is that the text was moderate in length and contained sentences which were not long and complicated.

As for type of sentences, more than 80% of the text was made up of simple sentences, while compound and complex sentences constituted 6.9% and 5.5% of the sentences in the text respectively. The high percentage of simple sentences thus made the story suitable for beginning ESL learners. This reflected that the selected text was suitable for the target learners. Table 3.6 depicts the factors in the readability of the text.

Table 3.6: Factors in the readability of the text

Item	Weighting
Lexical density	46.87 %
Density of unknown words	1:36
Sentence types Simple sentences Compound sentences Complex sentences	 80% 6.9% 5.5%

A fair amount of attention was given to the suitability of the text so that it formed comprehensible input for the learners; the plausible incomprehensibility factor would then be the target words in the text. Therefore, the study would be able to control the text in such a manner where the target words which were unfamiliar for the learners were the focus of the interaction. Table 3.7 provides a summary of how the input was selected.

Table 3.7: Summary of the criteria used for selection of input – the text

Input	Criteria	Method of Selection
	Expert agreement	Based on BEL 120 syllabus
Online narrative text	Suitability	Surface structure
Scary Night	Readability	Lexical density
		Density of unknown words

3.8.1.3 Target words

This process is discussed at two stages. The first stage was how the target words were selected and next was how these target words were glossed. The study heeded the call of Schmitt (2010) and Yoshii (2006) that working with only one word class would not confound the results of a study. Therefore, only verbs were used as target words and glossed. Thirteen verbs were selected. They were initially selected by two senior lecturers who taught semester one learners who agreed that the words were words that the learners were not familiar with. Then, these words were subjected to two other means of establishing that the words were not familiar to the sample of learners in the study.

The text with the target words was given to a similar group of learners as those in the experiment. It was found that the target words were also not familiar to these learners. Another approach to determine the target words was through their frequency as advocated by Schmitt (2010).

Using the British National (BNC) BNC 20,000 as a reference for the frequency of words, the profile revealed that the target words were from different frequency levels. The BNC-20 was used as the frequency information as it was a better representation of current English as found in other word lists (Schmitt, 2010). The BNC – 20 can be accessed in the Internet at http://www.lextutor.ca/vp/bnc/. The target words and their frequency bands range are shown in Appendix E.

A Frequency Band gives an indication of how frequent a word is used in the language. Words in the higher bands indicate that the words do not occur frequently, while the lower bands indicate otherwise. In this study, it can be observed that the frequency bands for the targeted word ranged from Bands 3 to 9. The Frequency band also showed that the words were not clustered in any particular band and it was dispersed, thus the target words were better represented in the text. Table 3.8 provides a summary of how the target words in the text were selected.

Table 3.8: Summary of selection criteria for target words

Input which had been modified	Selection criteria
Target verbs which have been glossed	Expert agreement
	Testing of familiarity of words with a similar group of learners as those in the study
	The British National Corpus (Frequency Band)

3.8.1.4 The gloss

The target verbs were glossed in four different formats. They were word definition in *Bahasa Melayu* in L1 (WBM) and English L2 (WEN); sentence level definition in *Bahasa Melayu* (SBM) and English (SEN).

The word definitions in Malay were as given in Hornby (2000) Learner's English-Malay dictionary, while the English word definitions were provided by Hornby (2005), Oxford Dictionary. Dictionaries were used as a source for word meanings as Schmitt (2010) contends that they are good sources as lexicographers have taken great measures to make the meanings as easy and transparent as possible.

The design of the gloss kept to Robb's (1999) taxonomy of gloss design for L2 reading. (Please see Appendix F for the taxonomy). Robb's taxonomy included authorship of glosses, their purpose, focus form and language (L1/L2) that was used for the gloss. The design of the gloss in this study was guided by this taxonomy. Although Robb stated that the taxonomy was not definite, it did provide a useful and starting framework for the design of a gloss.

Next is to explain the framework as used in this study. The gloss was authored by the researcher, and it was presented to the learners when they clicked on a word which they were not sure of or wanted more information on while reading the online text. The function of the gloss was to clarify lexical information. Its focus was textual and the languages used were English and *Bahasa Melayu*. The gloss in this study took on only verbal forms with no multimedia or non-textual elements.

The learners were assigned into different gloss conditions which provided word meanings at word and sentence levels. The gloss was definitional when it provided the learners with word meanings in either the L1 or L2 and it was elaborative when it contextualized the meanings of the words at sentence level.

The opportunities for obtaining modified input were made possible when learners clicked on the highlighted target words, hence presenting the learners with the different types of meaning at word or sentence level and in either the L1 or L2.

The clicking is then operationalised as the precursor to the interactions with the modified input. The clicks signal the start of the interaction with the modified input. The sequence of the learners' interactions with the gloss is shown in the Figure 3.1

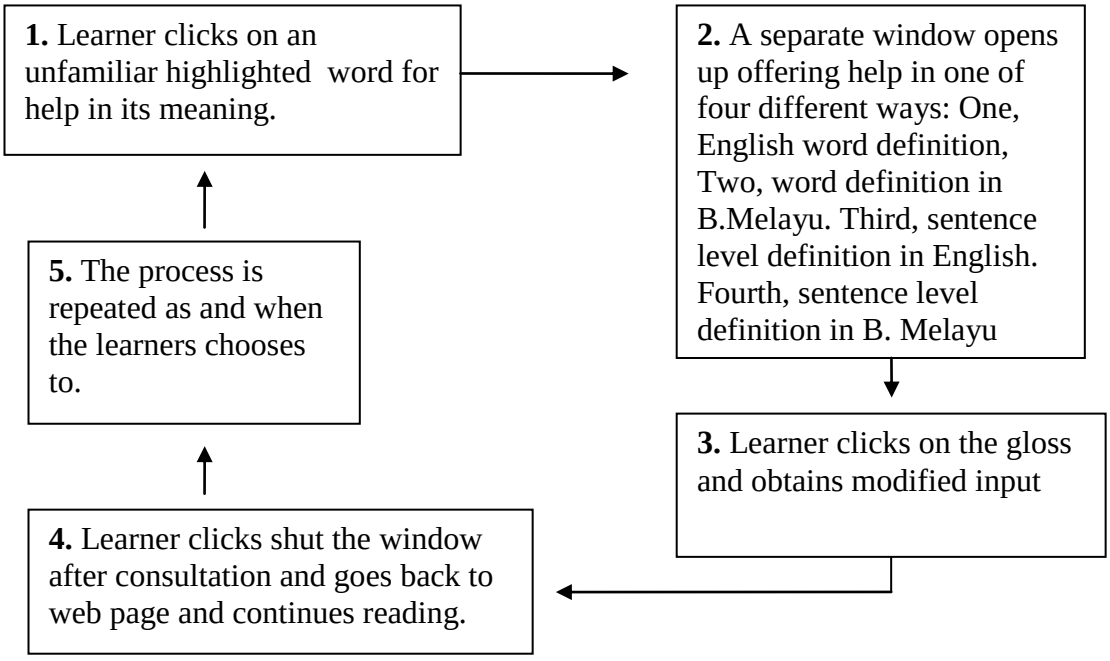


Figure 3.1: The sequence of the learners' interactions with the gloss

An example of this glossing process is shown below with the word "ponder"

The word in the text	I cannot pass the test. What do I do? Shall I keep studying? Can I take the test some other time? Shall I give up? I am **pondering *** many things.
	The Gloss for the word ponder
The Gloss	**Ponder**
Definitive	**Word meaning in English**: To think carefully about something, especially for a noticeable length of time
Contextual	**Sentence level meaning in English**: You need to <u>ponder</u> on what is the next step you should take in your life
Definitive	**Word meaning in BM**: *Memikirkan sesuatu dengan teliti dan ia mengambil masa.*
Contextual	**Sentence level meaning in BM**: *Selepas trajedi bot karam itu, Ahmad telah mula <u>memikirkan</u> soal hidup dan mati.*

3.8.1.5 The tracking device

It is not uncommon for research in CALL to use some kind of tracking device to track the pathways taken by learners with computers or recording their online behaviour on the Internet. In this research, a tracking programme was built into the online reading text; in this way the programme tracked the learners' interaction with the gloss.

The programme documented the moves by recording the types of gloss consulted and the frequencies of these look-ups. In short, the tracking device provided the empirical means to link possible look-up behaviour with vocabulary development. The tracking software was designed into the website to track all learners "moves" when they read the webpage.

The technical details of the tracking device are described here. The reading text was a web based application with a tracking device which was hosted at www.mohdalionline.com. Compared to normal web sites which only displayed static contents, this website (or web application, to be more specific) could process data and displayed the output to the programmer. To implement this web application, two important web application development tools had been used. The two tools are described below:

- Hypertext Pre-processor (PHP): this was the programming language or scripting language which was used to write codes which carry out specific tasks in the software, such as validating user credentials when the user logged-in and tracked the behaviour of learners when they surfed the main article.
- MYSQL: this tool was used to develop the database which acted as the software's back end component to store all the required information, such as learners' records, and the tracking results.

This website helped to track down the online behaviour of every learner who read the text by keeping track of the textual gloss consultations (word/sentence level meanings and language of the gloss). The website could be accessed at two levels:

(a) User Level: firstly, each learner accessed the website and registered himself/herself into the system and logged-in to be directed to the text. While reading the article, the learner could click on highlighted words to view their vocabulary information.

At the end of the session, learners clicked "Logout" to terminate their session.

(b) Administrator Level: The system administrator logged into the system at the end of the session to view the tracking results presented in tabular format, learner by learner or by groups (proficiency levels).

The tracker data can be analysed with the vocabulary test scores and the learners' proficiency levels. Thus, a more complete and accurate picture of the look-up behaviour of the learners could be obtained instead of just the vocabulary test scores (Hayden, 1997). The vocabulary tests are described next.

3.8.2 The Instruments: The Vocabulary tests

For this research three types of tests (pre-test, immediate post-test and delayed post-test) were deployed for different purposes. Cheng and Good (2009) describe the purpose of a pre-test is to assess existing knowledge. In this study it was to measure how many targeted words the learners knew before reading the online text. An immediate post-test was conducted to find out the effects of the treatment and in this case it was to find out if reading the online text with and without the gloss led to vocabulary expansion.

Finally, a delayed vocabulary test was to test if the learners had sustained the vocabulary knowledge over a period of time, which can indicate vocabulary learning. The lapse of three weeks was a reasonable time frame (Schmitt, 2010) to conduct a delayed test as the process of maturation may not have taken place and the test effects of the immediate post-test would have sufficiently eroded.

The purpose of the vocabulary tests in this research was to show evidence that there had been vocabulary gains or otherwise as a result of the learners' interactions with the gloss. At this juncture, it is important to bear in mind how the vocabulary construct was operationalised. It was imperative to make a distinction between the types of vocabulary knowledge tested as it would then fit into the types of vocabulary tests developed for this research. The vocabulary gain was measured by different tests to capture the different types of vocabulary knowledge, which were receptive and production. The tests were also repeated at set time intervals to measure whether vocabulary knowledge was maintained.

The tests designed in this study were to measure the learners' perceived, receptive and productive vocabulary knowledge. Therefore, a receptive test was used to find out if the learners could recall the meaning for the word form, while a productive test measured if the learners could produce the appropriate word form to express the meaning (Nation, 2001).

3.8.2.1 Pre-test: Measuring baseline vocabulary knowledge of the target words

There are various methods of measuring existing vocabulary knowledge. One of the methods is to use elicitation and translation methods. In his study Yoshii (2006) had used what he termed as "definition-supply" test where learners were asked if they knew a word, and if they knew it, they were to provide a definition of it in the L2.

The difficulty with this type of measurement is that it takes up time for learners to complete the test. In addition, the scoring of the word definition may pose problems. To counter this, Yoshii developed a scoring guide to ensure that scoring of the definitions in the L2 is within the parameters of the guide. This present study did not utilize the definition-supply test; instead it employed a self-rate vocabulary knowledge test.

This study acknowledges the use of translation in measuring vocabulary as it allows "learners to respond to vocabulary items in a way that does not draw on second language knowledge which is not directly relevant to what is being tested" (Nation, 2001, p. 351). However, it could also mean that learners who are more proficient than their peers would be able to express the meaning of the target word better compared to a learners who are less proficient in L1. Therefore, to counter this effect, this current study used a scale-type of vocabulary measurement where learners were not required to give the meanings of the words in L1 or L2; instead they were asked to self-rate their knowledge of the words.

The scale was first used by Horst and Meara (1999). The scale that was employed is given below:

0 = I definitely don't know what this word means

1= I am not really sure what this word means

2=I think I know what this word means

3=I definitely know what this word means

The learners in this study were asked to rate the words in the text according to the scale rather than demonstrate knowledge of the words. As with Horst and Meara's (1999) study, the strategy was to keep the learners' awareness of the target words at a minimum to avoid cross-test effects. The other reason for this kind of test is that the test is able to measure vocabulary knowledge of the learners at three separate points in

time of the study; vis-à-vis pre-experiment, immediately after experiment and delayed test. This test is designed for longitudinal studies, therefore it was appropriate for it to be used as the element of time interval was a factor in this study, that is, vocabulary knowledge of the learners were measured over a period of three weeks. Besides the useful information from this self-rate vocabulary knowledge test, it is easy to take with learners ticking the appropriate number and it does not take too much of the learners' time. In short it is easy to administer and mark.

To further explain this measure, this self-rate test was divided into Self-report categories, possible scores or points and the meanings of the scores. These criteria were adapted from Paribakht and Wesche (1997) Vocabulary Knowledge Score. The self-report test with its categories is shown in Table 3.9.

Table 3.9: Self-report categories, meanings and points

Self-report categories	Meanings of self-report categories	Points
0	I definitely don't know the meaning of the word	0
1	I am not really sure what this word means	1
2	I think I know what this word means	2
3	I definitely know the meaning of the word	3

In order to derive a single score for the vocabulary knowledge test, the following Computational method was used. The highest possible score for the test was 39 points, which is 13 words X 3. Following this the possible scores achieved by the learners were divided by three to obtain the single vocabulary knowledge test score. A single test score was necessary to compare the scores of the pre, immediate and delayed vocabulary tests.

The pre-test for the study included the 13 target words and the purpose of the test was to measure the learners' existing knowledge of the words. The test also included additional 10 distractors which were also verbs forms. The reason to have the distractors was not to sensitize the learners to the targeted items. This test acted as baseline vocabulary knowledge of the learners. This data were then compared to the same immediate and delayed vocabulary test scores. (Please see Appendix G for pre-test).

The self-rate test is really the simplest possible format for testing vocabulary knowledge. It is simple in the sense that it takes away the demands of proving that test takers know the words as in translation type of tests. However, the downside to this is that the validity of the test-takers in reporting what they know is difficult to gauge. Or the test-takers can have a different idea of what it is to know a word from the researcher.

3.8.2.2 The Post-tests

Two sets of post-tests were given to the learners. They were the immediate and delayed post-tests. They were the same but the items were scrambled. The post-tests consisted of three different tests. First was the vocabulary knowledge test where the learners were asked to self-rate the words according to the set categories. The second set of tests was the production and the word receptive tests.

Only post-tests were administered for the receptive and productive knowledge because it was to measure the learning after the treatment. In other words the researcher wished to observe the effect of the gloss use through the immediate test, after the experiment. These tests were not administered as pre-tests because it has been already established that there were benefits of using gloss in vocabulary learning (please see section 1.4) therefore it did make sense to measure the gains after the gloss use.

Further, if the tests were given as pre-tests, the learners would have been alerted that the research was interested in the target vocabulary, and would have made attempts at learning the vocabulary. The other effect would of course be that the learners would be attentive to the target words as they appear in three pre-tests.

It was necessary to place a delayed test in the experiment as vocabulary knowledge gained will usually be lost over time and a delayed post-test would show learning (Schmitt, 2010). The delayed test was conducted three weeks after the treatment. This delay of three weeks was considered as fair amount of time lapse between the treatment and the test. According to Schmitt (2010), "a delayed post-test of three weeks should be indicative of learning which is stable and durable" (p. 157). The post-tests are described next.

(a) Productive test: Gap-fill test

The productive test was in the format of a form-recall test (Laufer & Nation, 1999; Laufer & Goldstein, 2004). In particular, the test format can be termed as controlled productive vocabulary test (Meara & Fitzpatrick, 2000). In the test, the item format was a defining sentence context with a blank for the learners to fill in the correct word.

For this test, learners were asked to fill in the blanks in the sentences with the correct words from reading the online text. (Please see Appendix H for word productive test). The initial letter(s) of the word was/were provided to help the learners provide the targeted words instead of using other words. This was necessary so that they would not derive other synonymous words (Schmitt, 2010). In other words, this gap-fill test was a test which the researcher could control, at least to some degree, on the words that the learners will produce (Milton, 2009).

In this case, the test could measure the productive vocabulary knowledge of the learners. The items in this test were developed at sentence level, thus provided some form of context to the learners. The contexts were slightly different from the contexts found in the text. Nonetheless, the essential meanings of the words were still retained in the limited contexts of the productive test. This was deliberate as to assess the productive knowledge of the learners. The performance of learners in this test would show whether the learners' knowledge of a word was past the receptive stage and had moved towards productive knowledge (Schmitt, 2002).

(b) Word Receptive Test

The receptive test was in the form of multiple choice questions with four distractors. In this word receptive test, the learners had to select the meaning of the words. All the definitions were in English. The test was in the format of isolated words being matched to their meanings (Nation, 2001). (Please see Appendix I for word receptive test). Several factors were considered in the construction of the detractors. Firstly, the detractors were also in the verb form. Secondly, they were consistent in length and finally all the distracters were plausible answers.

Both these tests were administered after the treatment. There were no pre-tests for productive and receptive tests. As mentioned earlier, the reasons for this are two-fold. Firstly, the pre-test would probably sensitize the learners to the words tested which may affect the test scores as learners are able to figure the meanings of the words and secondly, the nature of these two tests was to measure the performance of the learners in the tests after using the gloss.

3.8.2.3 Scoring of the Receptive and Productive tests

All the items in the post-tests carried 1 mark each for a correct answer. As for the productive test, marks were awarded even if the word had inappropriate inflections, for example, "grins" instead of "grin". However, no marks were given for misspelt words. The test results were considered as continuous scores.

3.8.2.4 Reliability

As with any tests, reliability and validity issues had to be taken into account. The researcher assumed that the tests had already undergone some validation process where reliability concerns had been addressed by Yoshi (2006) is his research. Nonetheless, it was important to carry out the present study's reliability and validity checks as it was carried out in a different context (Rudestam & Newton, 2001). Reliability was proven by using Cronbach's alpha and the test results were 0.72 for the productive test and 0.78 for the receptive test.

On top of this, the multi-choice format of the receptive vocabulary test is considered credible in measuring vocabulary. This is stated by Daller, Milton and Treffers-Daller (2007) as "multiple-choice and forced answer test generally seem to have very good reliability and part of the credibility attached to tests of receptive vocabulary knowledge is that they give very reliable results" (p. 16). The other factors which may affect the reliability of the tests as stated by Perry (2011) were also taken into account in this study. The first factor is the subjectivity of the scorers or raters of the tests. In this case, it is the researcher himself who is the rater and as the answers to the tests are straightforward in the sense that there is only one right answer, thus this subjectivity bias is removed. The second factor is the length of the vocabulary tests.

With 13 items in the tests, it is fair to say that the length of the tests is neither too long nor short. The other factor related to reliability is the item quality used in the tests. The items were not ambiguous as the items which were identified as unfamiliar to the participants had adhered to accepted procedures such as expert agreement and using another cohort of students with similar traits to the participants to identify the unfamiliar words in the text (explained in section 3.8.1.3).

3.8.2.5 Validity

As for validity issues, the researcher conceded that it was difficult to establish criterion validity as it was difficult to benchmark these tests with established tests as none existed as most vocabulary tests measured different aspects of word knowledge in different kinds of circumstances with different sets of learners. Schmitt (2010) attests to this as he writes, "the complex nature of vocabulary knowledge dictates that any particular test would be severely limited as a criterion measure" (p. 181). However, the tests used in this study, which were receptive and productive, were common forms of tests that were used in research to measure vocabulary learning. The closest benchmark of the tests used here were to the ones used in Yoshii's (2006), albeit changes to the pre-test.

Schmitt (2010) explains that content validity of vocabulary tests could be provided by specifying the following elements: the tests should measure only specific lexical items or targeted words. In this case, the targeted words were words which were unfamiliar to the students. (Please refer to section 3.8.13 for a description of how the target words were selected). The next point was that the targeted words belonged only to one word class, in this study, they were the verb forms.

Furthermore, the verbs in the online text ranged from band three to band nine in the frequency band, providing a range of words in different bands.

For construct validity, the word knowledge that was being addressed was the form-meaning link and the tests that were designed measured this form-meaning aspect in the form of the receptive and productive vocabulary knowledge. For face validity, both these tests were designed in the test format of multiple-choice questions for the receptive vocabulary knowledge, while productive vocabulary knowledge was in the format of gap-filling sentences.

Both the formats are relatively familiar test formats to the students with the multiple-choice questions being a common format for the students. The production test format may not be that common as the receptive test therefore instructions to the students were written clearly at the beginning of the test to ensure that students understood what were required of them in the test.

3.9 Data collection procedure

The study consisted of three separate stages: the pre-experiment, experiment and post-experiment. A summary of the three stages is provided in Figure 3.2.

Stage 1	-	Consent from learners
Pre-experiment	-	Pre-vocabulary knowledge test for all groups. Perceived vocabulary knowledge test 0
Stage 2	-	Learners stratified into low, mid and high proficiency levels
		Learners randomized into different gloss conditions
(Experiment)	-	Reading the online text by all groups

- Immediate post-tests after the treatment . Perceived vocabulary knowledge test 1, Production test 1, Receptive test 1

Stage 3 Administration of delayed post-tests. Perceived vocabulary knowledge test 2, Production test 2 and Receptive test 2 to all groups
(Post Experiment)

Figure 3.2: The stages of the study

3.9.1 Pre-experiment

Prior to the experiment, the researcher explained the nature of the study to the learners. They were told that it was an online reading task with a reading comprehension exercise after they had read the text. They were informed that the data collected would be for the research only. Furthermore, it was important to assure them that their performance in the study would not affect their grades for the English course that they were taking. Learners who agreed to take part in the study were asked to sign a consent form (Please see Appendix J). Once this was done, the learners were given the pre-test which measured their knowledge of the target vocabulary items. This pre-test was given a week before the treatment was administered to learners and was done during class time for both groups. The time gap of one week was necessary between the pre-test and the actual experiment to avoid sensitizing the learners to the target words.

3.9.2 The Experiment

In the second stage, four language laboratories at UiTM were used for the study. Each laboratory was able to accommodate 35 learners. The learners were randomly assigned to the different gloss condition groups.

The number of learners in each gloss condition is given in the following table.

Table 3.10: Gloss conditions and number of learners

Gloss Condition	Number of learners
W BM	22
SBM	23
W EN	24
SEN	30
Total	99

After this, each learner was assigned to a computer with Internet connection to access the online reading text. Once this was done a briefing session was given to the learners. They were told how to use the gloss, that is, by clicking on the highlighted words which would then open up the different types of glosses for the word clicked.

The learners were told to access the text at www.mohdalionline.com. Before they read the text, they were asked to fill in their details: UiTM student number, name and SPM English grade as requested by the computer programme. They were then directed to the webpage which contained the story "A Scary Night". After completing the online reading task, the learners were told to log off from the website. After this, they were given three immediate post-tests sequentially: the perceived vocabulary knowledge, production and receptive vocabulary tests. The tests were in the pencil-and-paper format.

3.9.3 Post Experiment

After three weeks the learners were given the delayed post-tests unexpectedly during their regular class time and this concluded the third stage of the study. The sequence of the delayed tests was the same as the immediate post-tests. Firstly, the perceived vocabulary knowledge, followed by the production and finally the word receptive test. The content of these delayed tests was the same as the immediate tests.

However, the order of the test items was different to avoid the testing effect, in which learners might have been "test wise" if the items were to remain in the same order. This concluded the third stage of the data collection procedure. Table 3.9 depicts the sequence and types of tests used in the study.

Table 3.11: The sequence and types of tests for the study

Group	**Pre-test** (Administered 1 week before the treatment)	**Treatment**	**Immediate post-tests** (Administered immediately after the treatment)	**Delayed post-tests** (Administered 3 weeks after the treatment)
All gloss conditions	Perceived vocabulary knowledge pre-test 0	Read the online text with the glosses	-Perceived vocabulary knowledge pre-test 1 -Production test 1 -Receptive test 1	-Perceived vocabulary knowledge pre-test 2 -Production test 2 - Receptive test 2

3.10 Conclusion

Working onwards from the purpose and research questions of the study, this chapter went on to describe the final 99 participants from UiTM who took part in this research. The means to collect the data which were the online reading text (the input) and gloss which formed the basis of the computer-learner interactions were also detailed to show how data were elicited. On top of that, the vocabulary tests which consisted of pre- and post-test were described. They consisted of the perceived vocabulary knowledge, receptive and productive vocabulary tests.

CHAPTER 4

DATA ANALYSIS AND RESULTS

4.1 Introduction

The main aim of this study is to investigate how students use a computer-aided textual gloss to enhance their vocabulary knowledge. This is done through a quasi experimental set-up where students read an online reading text which was embedded with the textual gloss. The gloss was designed to give meanings of unfamiliar words in the text at word and sentence levels, in the students L1 (*Bahasa Melayu*) and L2 (English).

Before the thesis presents the findings of the experiment based on the research questions, it is best to state here that the unit of analysis of the study was on the learner-computer interactions. Next, this section gives a general description of the scope and different layers of the analysis. Analysis first begins with data on the clicking behaviour of the learners and whether these interactions do facilitate vocabulary development. After this was done, there was further examination of the effect of the specific type and language of gloss on the learners' vocabulary knowledge.

The analysis was also carried out on two levels. One was within-subject that is the performance of the learners in the different vocabulary knowledge tests. The other was between-subject where the performance of the learners was compared between the different proficiency levels of low, medium and high. A final total of 99 students took part in the experiment out of the original sample of 117 students as some of them did not complete the set of vocabulary tests, while some students' interactions were not fully tracked by the tracking device.

4.2 Clicking behavior

Research Question 1:

1. What is the clicking behaviour of the learners (a) in the different gloss conditions (b) of different proficiency levels?

4.2.1 Clicking behaviour in the different gloss conditions

A one way analysis of Variance (ANOVA) was conducted to assess whether there were differences between total number of clicks in the 4 gloss conditions. Preliminary assumption testing was conducted to check for normality and homogeneity of variance of the dependent variable. The results showed no departure from the two assumptions for all proficiency levels. Table 4.1 depicts the mean and standard deviation for all tests in four gloss conditions for total number of clicks. Figure 1 shows the mean and gloss conditions.

Table 4.1: Means and standard deviations for the number of clicks in each gloss condition

	N	Minimum	Maximum	Mean	Std. Deviation
W BM	22	2	24	13.4091	4.90538
S BM	23	0	26	12.2609	7.05960
W EN	24	0	24	8.4583	5.80089
S EN	30	1	24	9.0000	5.97697
Total	99	0	26	10.6061	6.25613

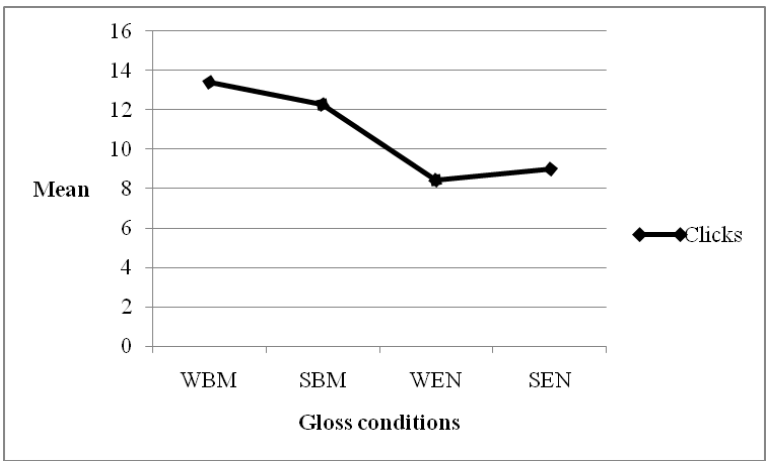


Figure 4.1: The mean number of clicks in each gloss condition

Table 4.2: ANOVA for number of clicks in four gloss conditions

	df	Mean Square	F	Sig.
Between Groups	3	141.308	3.935	.011
Within Groups	95	35.913		
Total	98			

The results of one way ANOVA for the vocabulary knowledge test are presented in Table 4.2. The ANOVA test revealed that there was a statistically significant difference in the mean total click numbers in the 4 gloss conditions F (3, 26) =3.93, p=0.011.

The effect size is 0.11 which is considered a large effect size according to Cohen (1998).

Table 4.3: Tukey's Post-Hoc comparisons results

	(I) Glossary Type	(J) Glossary Type	Mean Difference (I-J)	Std. Error	Sig.
Tukey HSD	W BM	SBM	1.14822	1.78713	.918
		WEN	4.95076*	1.76883	.031
		SEN	4.40909*	1.68211	.049

The Tukey's Honestly Statistically Differences (HSD) post-hoc multiple comparisons (Table 4.3) showed a statistically significant difference in the following pairs:

Word BM gloss (*M*=13.40, *SD* =4.90) and Sentence English gloss (*M*= 9.00, *SD* = 5.97) with a mean difference (*MD*=4.4) and the p-value of 0.049.

Word BM gloss (*M*=13.40, *SD* =4.90) and word English gloss (*M*= 8.45, SD = 5.80) with a large mean difference (*MD* =4.95) and the p-value of 0.031.

From the data, it can be seen that the learners in the word BM gloss condition had clicked on the target words the most. This was followed by the sentence BM gloss condition. The least number of clicks were made by the learners in the Word EN gloss condition. Comparatively, the difference between the word BM and word EN types of glosses was large. This meant that learners in the word BM clicked the most number of words, and the learners in the word EN gloss conditions clicked the least.

Data from research question 1 also suggest that those learners had the tendency to click more on the target words if the gloss was in the word BM condition. This can be interpreted as the learners found that it was helpful and clicked on the gloss to get the meaning of the target words. The other gloss condition where the learners clicked frequently was the sentence BM condition. As for the sentence BM group, again what can be highlighted is that the language of the gloss was still BM, the learners L1. Overall, it can be said then that glosses in L1 encouraged the learners to click more on the glosses.

The lowest number of clicks was recorded by the learners in the Word EN gloss condition. A reason for this could be that the learners in this gloss condition may not find the gloss which is configured in the word and English combination worth clicking as they may be perceived by them as not being helpful. This is the initial interpretation of the data at this stage based solely on the number of clicks.

4.2.2 Clicking behaviour of learners of different levels of proficiency

A one way analysis of Variance (ANOVA) was conducted to assess whether there were differences between the total number of clicks in the different proficiency levels. The dependent variable in each ANOVA was the total number of clicks and the independent variable was the proficiency levels. If significant results were found, then Tukey's post -hoc multiple comparisons were further computed to determine where the difference between the glosses condition was.

Preliminary assumption testing was conducted to check for normality and homogeneity of variance of the dependent variable. The results showed no serious departure from the two assumptions for all proficiency levels. Table 4.4 indicates the mean and standard deviation of the number of clicks at low, mid and high proficiency level. A closer inspection of the mean does show some differences, although insignificant, in the clicking behaviour of the learners in different proficiency levels. From the means, it can be observed that the mid proficiency group had the most number of clicks (M=11.70 ,SD=6.3) and high proficiency level had the least number of clicks (M=9.57 , SD=5.28). This is depicted in Figure 4.2, where it can be seen that the mid proficiency level learners had clicked the most number of target words in the text.

Table 4.4: Mean and standard deviation for number of clicks in each proficiency level

	N	Minimum	Maximum	Mean	Std. Deviation	Std Error
LOW	30	0	26	10.2333	7.01566	1.28088
MID	39	0	25	11.6923	6.32103	1.01218
HIGH	30	0	21	9.5667	5.27638	.96333
Total	99	0	26	10.6061	6.25613	.62876

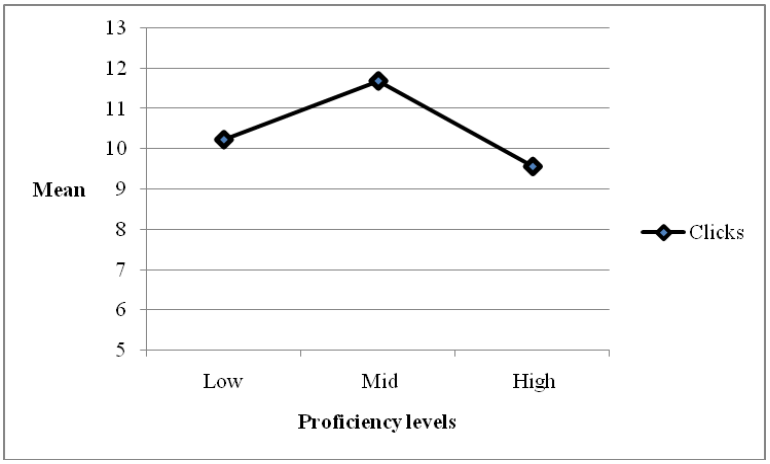


Figure 4.2: The mean number of clicks by the learners in different proficiency levels

The results of one way ANOVA are presented in Table 4.5. ANOVA and post-hoc comparisons both indicated that in low proficiency level there was no statistically significant difference between different proficiency levels in the number of clicks. F (2, 96) =1.056 P-value=0.352. This indicates that the clicking behaviour of the learners in the different proficiency levels was almost the same.

Table 4.5: ANOVA for number of clicks at different proficiency levels

	df	Mean Square	F	Sig.
Between Groups	2	41.298	1.056	.352
Within Groups	96	39.094		
Total	98			

From the data, initially it is revealed that there was no significant difference in the way the learners of different proficiency levels clicked the target words. However, a tighter analysis showed that the mid proficiency level learners clicked on more words as compared to the low and high proficiency learners. The data also revealed that the BM glosses were clicked more than English ones. In addition, word type of glosses was clicked more by the learners.

4.3 Interaction with glosses

Research question 2 (a):

Which specific type of modified input, that is, word or sentence and in learners' L1 or L2 in the interactions facilitates (a) perceived vocabulary knowledge (b) productive vocabulary knowledge (c) receptive vocabulary knowledge of learners with different language proficiency levels?

4.3.1 Perceived vocabulary knowledge

In this part, the researcher investigated the effect of language (BM and EN) and type of gloss (word and sentence). This study incorporated the between-subjects variables of language with two levels (BM and EN), the between-subject variable of type of gloss with two levels (word and sentence) and, the within-subject variable of time (pre-test, immediate post-test and delayed post-test). Hence a Mixed ANOVA was carried out to get a general pattern of the effects of the language (BM, EN) and type (word, sentence) of glosses and the changes that took place from the perceived vocabulary knowledge test at three different points in the research, pre, immediate and delayed.

4.3.1.1 Low proficiency level

The assumptions of normality and Sphericity were tested. The assumption of normality met The Mauchley's test for Sphericity which was significant ($p=0.003$), thus the assumption of Sphericity was not met and Greenhouse-Geisser

adjustment was used. Table 4.6 and Table 4.7 present the descriptive statistics for test scores for the different gloss conditions.

Table 4.6: Mean and standard deviation for each language in low proficiency group

	Language	Mean	Std. Deviation	N
Pre-test score	BM	2.6667	1.49691	13
	EN	1.5882	1.44111	17
Immediate test score	BM	6.9746	3.00093	13
	EN	5.6863	3.04702	17
Delayed test score	BM	6.1279	3.09923	13
	EN	4.9412	2.81917	17

Table 4.7: Mean and standard deviation for each gloss type in low proficiency group

	Type	Mean	Std. Deviation	N
Pre-test score	Word	1.9722	1.79482	12
	Sentence	2.1111	1.39560	18
Immediate test score	Word	6.4169	2.92320	12
	Sentence	6.1296	3.20159	18
Delayed test score	Word	3.8886	2.62625	12
	Sentence	6.5000	2.74219	18

Figure 4.3 shows that for both BM and EN glosses the test scores increased from pre-test to immediate test and then decreased from immediate test to delayed test. Since there was not a significant time*language interaction the lines on the profile plot were parallel. Figure 4.4 shows that for both word and sentence glosses the test scores increased from pre-test to immediate test. For subjects whose used word gloss the scores decreased from immediate test to delayed test while for sentence gloss group the scores increased from immediate test to delayed test.

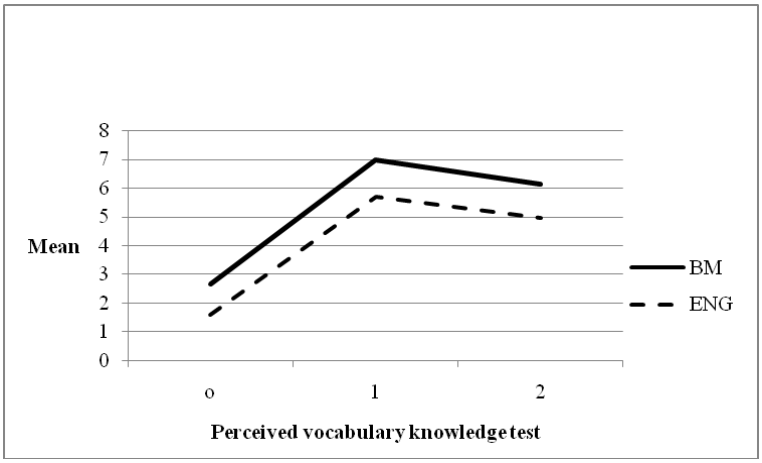


Figure 4.3: Language of glosses for perceived vocabulary knowledge tests in low proficiency level

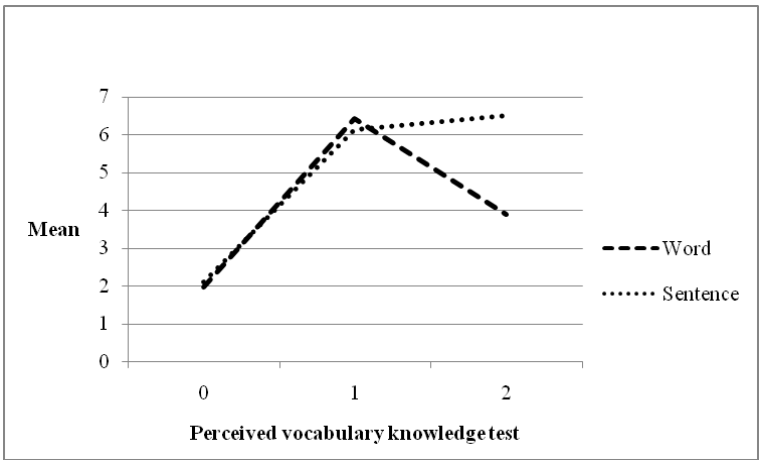


Figure 4.4: Type of glosses for perceived vocabulary knowledge tests in low proficiency level

(a)Test of between-subject effects

The results of tests for between-subject effects indicated that for subjects in low proficiency level neither the language of gloss $F(1, 26) = 2.149$, $P < 0.155$ $\eta^2 = 0.076$ nor the type of gloss $F(1, 26) = 2.225$, $P=0.148$, $\eta^2 = 0.079$ had a significant effect on vocabulary development. The language*type interaction was not significant.

It meant that there were no differences among the BM and English glosses, and between word and sentence glosses. The results of Mixed ANOVA are summarized in Table 4.8.

Table 4.8: Results of Mixed ANOVA in low proficiency level

Source	df	MSE	F	p-value	η^2
Time	1.455	192.743	47.632	.000**	.647
Time * Language	1.455	1.309	.324	.656	.012
Time * type	1.455	21.129	5.222	.017 *	.167
Time * Language*type	1.455	17.943	4.434	.029 *	.146
Error (within subjects)	37.832	4.046			
Language	1	23.933	2.149	.155	.076
type	1	24.773	2.225	.148	.079
Language*type	1	34.979	3.141	.088	.108
Error (between)	26	11.135			

(b)Test of within-subject effects

The main effect for the within-subject variable of time was significant F (1.445, 37.832) = 47.632, P = 0.000, η^2 = 0.647 indicating that there were differences among three test scores. The magnitude of the differences in the means was large (η^2 = .647). Comparing the obtained eta-squared value according to Cohen's (1988) criteria (0.01 = small effect, 0.06 = moderate effect, and 0.14 = large effect), showed it had a very large effect size. A large effect size suggested that the mean differences of the pre-test, immediate test and delayed test in low proficiency group were very large.

For the time*type interaction and Time * Language*type the results indicated that these interactions were statistically significant F (1.445, 37.832) = 5.222, P = .017, η^2=0.167 (for time*type interaction) and F (1.445, 37.832) = 4.434, P = .029, η^2=0.146 (for Time * Language*type interaction).

It was found before both type and language effects were not significant, however, the type main effect was qualified by a significant interaction between time and type.

As the interaction was statistically significant, the 'simple effects' that is, differences between means for one variable (time) at each particular level of the other variable (type) should be analysed according to Leech, Barrette & Morgan (2008). Since the time*type interaction was significant, the researcher followed the analysis by running an independent sample t-test in order to compare the word and sentence glosses in different test conditions (pre, immediate and delayed). The assumption of normality was checked and it was met. Levene's Test of equality of variance indicated that the variances of the all test scores for word and sentence groups were equal (p >0.05).

The results of independent sample t-test revealed that there was a statistically significant difference in the mean delayed test scores between word (Mean=3.89, SD=2.62) and sentence glosses (Mean=6.50, SD=2.74). t (28)=2.598, p=.015. There was no difference in the mean score and mean immediate test score for the word and sentence glosses.

4.3.1.2 Mid Proficiency

The assumptions of normality and Sphericity were tested. The assumption of normality was met. The Mauchley's test for Sphericity was not significant (p=0.734), thus the assumption of Sphericity was met. Tables 4.9 and 4.10 present the descriptive statistics for test scores for different gloss conditions in mid proficiency level.

Table 4.9: Mean and standard deviation for each language in mid proficiency level

	Language	Mean	Std. Deviation	N
Pre-test score	BM	2.7037	1.17651	18
	EN	4.5714	1.80476	21
Immediate test score	BM	8.3891	3.18099	18
	EN	7.9335	2.85037	21
Delayed test score	BM	7.8704	2.88140	18
	EN	7.7937	2.33447	21

Table 4.10: Mean and standard deviation for each gloss type in mid proficiency level

	Type	Mean	Std. Deviation	N
Pre-test Score	Word	3.4035	1.24513	19
	Sentence	4.0000	2.18982	20
Immediate Test Score	Word	8.5440	3.41752	19
	Sentence	7.7546	2.49666	20
Delayed Test Score	Word	7.7719	2.57013	19
	Sentence	7.8833	2.62751	20

The visual inspection of the profile plot in Figure 4.5 reveals that for both BM and EN glosses the test scores increased from pre-test to immediate test and then decreased from immediate test to delayed test. However, the rate of increase and decrease was different and the interaction effect was significant. This meant that BM glosses had helped learners obtain better scores in the immediate test but it was not maintained over time. On the other hand, it could also mean that the students simply knew less words at the pre-test, hence the better scores in the immediate test.

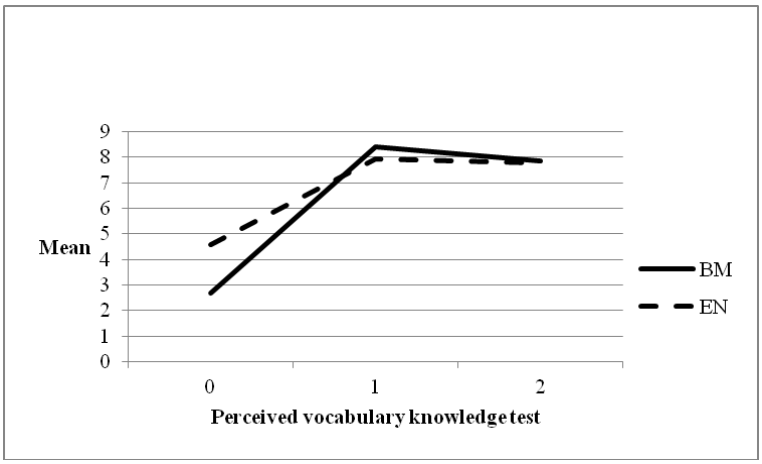


Figure 4.5: Language of glosses for perceived vocabulary knowledge test scores in mid proficiency level

A close inspection of Figure 4.6 shows that learners gained from both word and sentence type of glosses. Gains were also seen in the word type of glosses in the immediate test, however, learners who had interacted with the sentence type glosses showed an increase for the immediate to delayed tests.

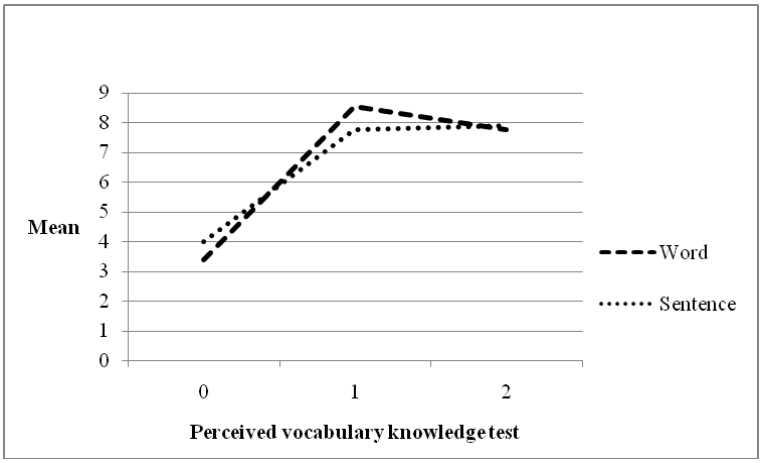


Figure 4.6 Type of glosses for perceived vocabulary knowledge tests in mid proficiency level

(a)Test of between-subject effects

The results of tests for between-subject effects indicated that for subjects in mid proficiency level neither the language of gloss F (1, 34) = 0.349, P < 0.535 η^2 = 0.011 nor the type of gloss F (1, 34) = 0.056, P=0.814, η^2 = 0.002 had a significant effect on vocabulary knowledge development. The language*type interaction was not significant. It meant that the there were no differences among the BM and English, and between word and sentence glosses. The results of Mixed ANOVA are summarized in Table 4.11.

Table 4.11: Results of Mixed ANOVA in mid proficiency level

Source	df	MSE	F	p-value	η^2
Time	2	236.312	56.965	.000**	.626
Time * Language	2	15.831	3.816	.027 *	.101
Time * type	2	4.026	.970	.384	.028
Time * Language*type	2	1.901	.458	.634	.013
Error (within subjects)	68	4.148			
Language	1	3.964	.394	.535	.011
type	1	.568	.056	.814	.002
Language*type	1	9.390	.933	.341	.027
Error (between)	34	10.068			

(b)Test of within-subject effects

The main effect for the within-subject variable of time was significant F (2, 68) = 56.965, P = 0.000, η^2 = 0.626 indicating that there were differences among three test scores with a very large effect size. For the interaction effects only time*language effect was statistically significant F (2, 68) = 3.816, P = .027, η^2=0.101.

The results of test of between-subject effects showed both type and language effect was not significant. However, since the time*language interaction was significant the researcher followed-up the analysis by running an independent sample t-test in order to compare the BM and English glosses in different test conditions (pre, immediate and delayed).

For the t-test, the assumption of normality was met. The test of equality of variance indicated that the variances of the all test scores for word and sentence groups were equal *(p* >0.05). The results of independent sample t-test revealed that only there was a statistically significant difference in the mean scores between BM and English glosses *t* (37) =3.756, *p*=.001. There was no difference in the mean immediate test score and mean delayed test score for the word and sentence glosses since the results didn't show any differences.

4.3.1.3 High Proficiency

The assumptions of normality and Sphericity were tested. The assumption of normality met The Mauchley's test for Sphericity was significant (p=0.002), thus the assumption of Sphericity was not met and Greenhouse-Geisser adjustment was used. Tables 4.12 and 4.13 present the descriptive statistics for test scores for different gloss conditions for high proficiency level learners.

Table 4.12: Mean and standard deviation for each language in high proficiency level

	Language	Mean	Std. Deviation	N
Pre-test score	BM	5.0238	2.20901	14
	EN	5.9167	2.84279	16
Immediate test score	BM	9.4762	3.05145	14
	EN	10.5417	1.80483	16
Delayed test score	BM	8.5476	3.18546	14
	EN	10.6042	1.35657	16

Table 4.13: Mean and standard deviation for each gloss type in high proficiency level

	Type	Mean	Std. Deviation	N
Pre-test score	Word	5.4222	2.62306	15
	Sentence	5.5778	2.59262	15
Immediate test score	Word	10.6222	2.52249	15
	Sentence	9.4667	2.37981	15
Delayed test score	Word	9.3111	3.05626	15
	Sentence	9.9778	2.01765	15

The visual inspection of the profile plot in Figure 4.7 reveals that for both BM and English glosses, the test scores increased from the pre-test to immediate test. For subjects who used English glosses the scores had a small increase from immediate test to delayed test while for BM glosses group the scores decreased from immediate test to delayed test.

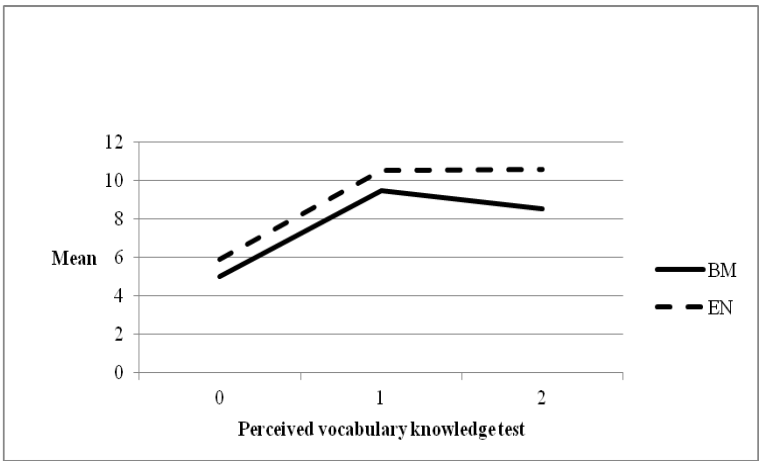


Figure 4.7: Language glosses for perceived vocabulary knowledge test scores in high proficiency level

Figure 4.8 shows that for both word and sentence gloss the test scores increased from pre-test to immediate test. For learners who had interactions with the word gloss, the scores decreased from immediate test to delayed test while interactions with sentence gloss group the scores increased from immediate test to delayed test.

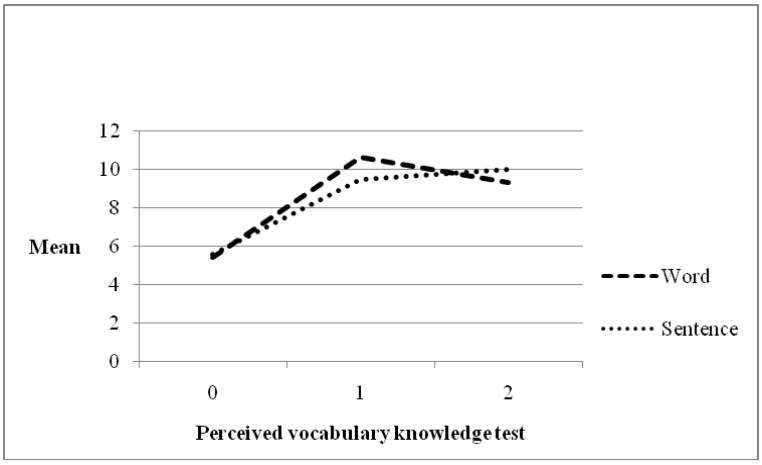


Figure 4.8: Type of glosses for perceived vocabulary knowledge test scores in high proficiency level

(a)Test of between-subject effects

The results of tests for between-subject effects indicated that for subjects in high proficiency level neither the language of gloss F (1, 26) = 3.099, P =.090 η^2 = 0.107 nor the type of gloss F (1, 26) = 0.014, P=0.906, η^2 = 0.001 had a significant effect on vocabulary development. The language*type interaction was not significant. It meant that there were no differences among the BM and English glosses, and between word and sentence glosses. The results of Mixed ANOVA are summarized in Table 4.14.

Table 4.14: Results of Mixed ANOVA in high proficiency level

Source	df	MSE	F	p-value	η^2
Time	1.441	260.110	62.338	.000**	.706
Time * Language	1.441	4.087	.980	.360	.036
Time * type	1.441	9.959	2.387	.120	.084
Time * Language*type	1.441	3.806	.912	.381	.034
Error (within subjects)	37.460	4.173			
Language	1	40.119	3.099	.090	.107
type	1	.183	.014	.906	.001
Language*type	1	2.187	.169	.684	.006
Error (between)	26	12.945			

(b)Test of within-subject effects

The main effect for the within-subject variable of time was significant F (1.441, 37.46) = 62.338, P = 0.000, η^2 = 0.706 indicating that there were differences among three test scores with a very large effect size.

4.3.2 Productive vocabulary knowledge

A mixed between-within ANOVA with two between-subject (language and type of gloss) and one within-subject variable (word production tests 1 and 2) was conducted to investigate the effect of language and type of gloss on productive vocabulary knowledge. This study incorporated the between-subject variable of language with two levels (BM and EN), the between-subject variable of type of gloss with two levels (word and sentence) and, the within-subject variable of time (word production test 1 and word production test 2). The Mixed ANOVA tested whether there were main effects for each of the independent variables and whether the interactions between the variables were significant. The following are the results for the different proficiency levels.

4.3.2.1 Low Proficiency

Preliminary assumption testing was conducted to check for normality and homogeneity of variance of the mean test scores. Initial results showed that the two assumptions were fulfilled. Since the within-subject variable of time had only two levels (test 1and test 2) so it was not necessarily to check the assumption of Sphericity for all three proficiency levels. Table 4.15 and Table 4.16 present the descriptive statistics for test scores for different gloss conditions.

Table 4.15: Mean and standard deviation for each language in low proficiency

	Language	Mean	Std. Deviation	N
VOCABULARY PRODUCTIVE TEST 1	BM	12.15	1.144	13
	EN	11.29	1.160	17
VOCABULARY PRODUCTIVE TEST 2	BM	11.15	2.267	13
	EN	10.18	2.921	17

Table 4.16: Mean and standard deviation for each gloss type in low proficiency level

	Type of gloss	Mean	Std. Deviation	N
VOCABULARY PRODUCTIVE TEST 1	Word	11.17	.937	12
	Sentence	12.00	1.283	18
VOCABULARY PRODUCTIVE TEST 2	Word	10.17	2.329	12
	Sentence	10.89	2.888	18

Figure 4.9 shows that the learners obtained higher scores when interacting with the BM glosses than with the English glosses across both test 1 and test 2, although these differences were not significant. In both languages (BM/EN), the test scores decreased from test 1 to test 2.

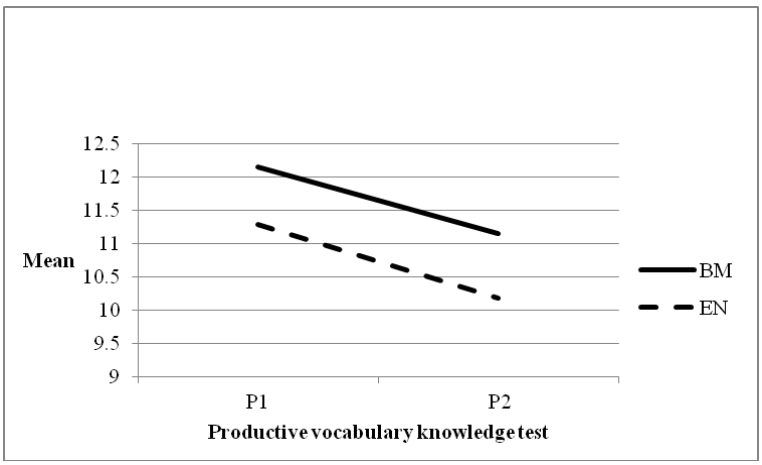


Figure 4.9: Language glosses for productive knowledge test scores in low proficiency level

Figure 4.10 shows that the learners obtained higher scores when interacting with sentence glosses than under the word glosses across both test 1 and test 2, although this difference was not significant (because the main effect of type was not significant). On the loss of productive vocabulary knowledge, it can be observed that in both gloss types (word and sentence) the test scores decreased from test 1 to test 2.

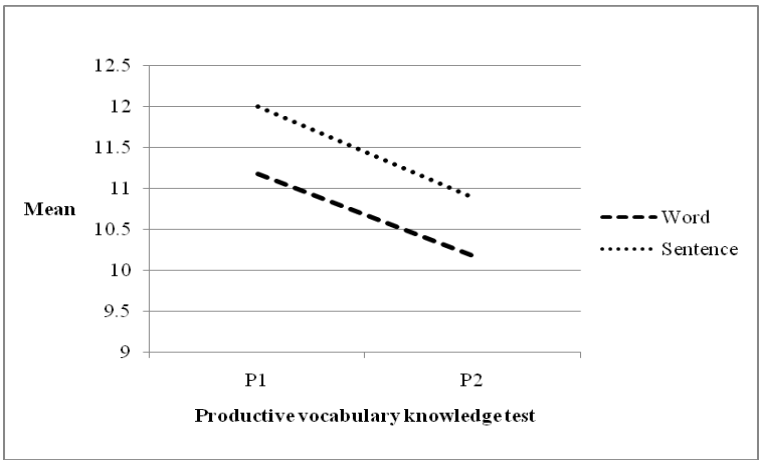


Figure 4.10: Type of glosses for productive vocabulary knowledge test scores in low proficiency level

(a)Test of within-subject effects

The main effect for the within-subject variable of time was not significant F (1, 26) =4.1 P=0.053, η^2 = 0.136 indicating that there were no large differences among two test scores. Before the main effects can be examined, there is a need to assess the interaction effect. The results of Mixed ANOVA are summarized in Table 4.14. As can be seen in Table 4.17 all two-way and three-way interactions (time*type, time*language and time*type*language) were insignificant.

Table 4.17: Results of Mixed ANOVA in low proficiency level

Source	df	MSE	F	p-value	η^2
Time	1	15.038	4.100	0.053	.0136
Time * Language	1	.006	.002	.968	.000
Time * type	1	.015	.004	.950	.000
Time * Language*type	1	.495	.135	.716	.005
Error (within subjects)	26	3.667			
Language	1	13.773	2.800	.106	.097
type	1	11.548	2.348	.138	.083
Language*type	1	.338	.069	.795	.003
Error (between)	26	4.918			

(b)Test of between-subject effects

The results of tests of between-subject effects indicated that for subjects in low proficiency level neither the language of gloss F (1, 26) = 2.80, P=0.106 η^2 = 0.097 nor the type of gloss F (1, 26) = 2.348, P=0.138, η^2 = 0.083 had a significant effect on vocabulary acquisition. The language*type interaction was not significant. There was no significant difference in the test scores between the two languages (BM and EN) and between two types of gloss (word and sentence).

4.3.2.2 Mid Proficiency

Tables 4.18 and 4.19 present the descriptive statistics for test scores for different gloss conditions.

Table 4.18: Mean and standard deviation for each language in mid proficiency level

	Language	Mean	Std. Deviation	N
VOCABULARY PRODUCTIVE TEST 1	BM	12.44	.705	18
	EN	12.00	1.049	21
VOCABULARY PRODUCTIVE TEST 2	BM	10.67	1.847	18
	EN	10.43	2.087	21

Table 4.19: Mean and standard deviation for each type of gloss in mid proficiency level

	Type of gloss	Mean	Std. Deviation	N
VOCABULARY PRODUCTIVE TEST 1	Word	12.05	1.079	19
	Sentence	12.35	0.745	20
VOCABULARY PRODUCTIVE TEST 2	Word	10.32	2.311	19
	Sentence	10.75	1.585	20

To aid the interpretation of the interaction effect, it would be useful to examine the graphs presented in Figures 4.11 and 4.12. Figure 4.11 shows that the learners obtained higher scores under the BM gloss condition than under the EN gloss condition in test 1 (although these differences were not significant and learners obtained almost the same mean scores in word productive test 2. In both gloss languages (BM/EN), the test scores decreased from test 1 to test 2.

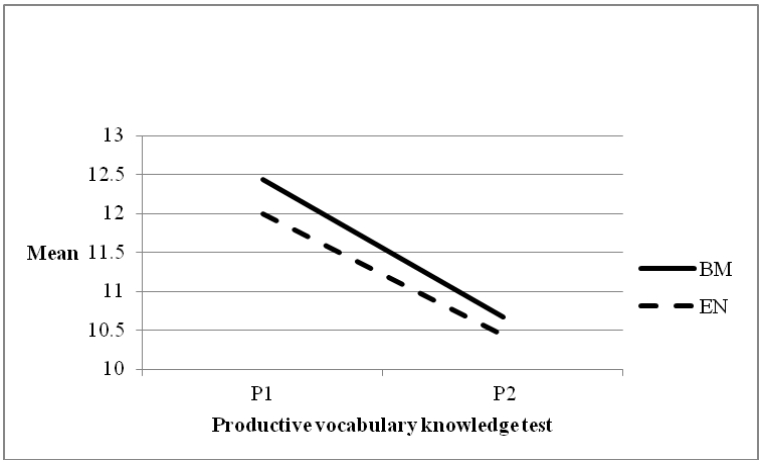


Figure 4.11: Language of glosses for productive knowledge test scores in mid proficiency level

Figure 4.12 shows that the learners obtained higher scores using sentence glosses than under word glosses across both test 1 and test 2, although these differences were not statistically significant (because the main effect of type was not significant). In both gloss types (word /sentence), the test scores decreased from test 1 to test 2.

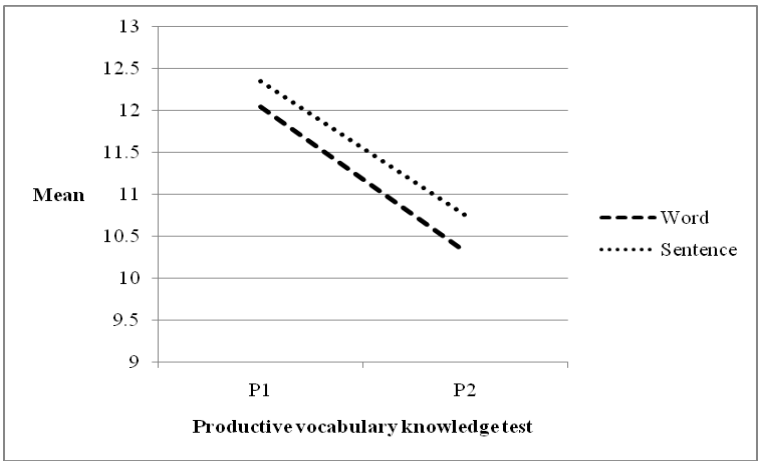


Figure 4.12: Type of glosses for productive knowledge test scores in mid proficiency level

(a)Test of within-subject effects

The main effect for the within-subject variable of time was significant, F (1, 35) =26.741, P=.000, η^2 = 0.433 indicating that there were significant differences among mean tests score between test 1 and test 2 with a large effect size. The results of Mixed ANOVA are summarized in Table 4.20. As can be seen in Table 4.20 all two-way and three way interactions (time*type, time*language and time*type*language) were insignificant.

Table 4.20: Results of Mixed ANOVA in mid proficiency level

Source	df	MSE	F	p-value	η^2
Time	1	54.641	26.741	.000**	.433
Time * Language	1	.186	.091	.765	.003
Time * type	1	.055	.027	.870	.001
Time * Language*type	1	.525	.257	.615	.007
Error (within subjects)	35	2.043			
Language	1	2.448	.860	.360	.024
type	1	2.448	.860	.360	.024
Language*type	1	1.157	.407	.528	.011
Error (between)	35	2.845			

(b)Test of between-subject effects

The results of tests for between-subject effects indicated that for subjects in mid proficiency level neither the language of gloss F (1, 35) = 0.860, P=0.360 η^2 = 0.024 nor the type of gloss F (1, 35) =2.448, P=0.360, η^2 = 0.024 had a significant effect on vocabulary development. The language*type interaction was not significant. There was no significant difference in the test scores for the two types of gloss (word and sentence) and the two languages (BM and EN).

4.3.3.3 High Proficiency

Table 4.21 and Table 4.22 present the descriptive statistics for test scores for different gloss conditions.

Table 4.21: Mean and standard deviation for each language gloss in high proficiency level

	Language	Mean	Std. Deviation	N
VOCABULARY PRODUCTIVE TEST 1	BM	12.50	.650	14
	EN	12.88	.342	16
VOCABULARY PRODUCTIVE TEST 2	BM	11.93	1.207	14
	EN	12.44	1.342	16

Table 4.22: Mean and standard deviation for each type of gloss in high proficiency level

	Type of gloss	Mean	Std. Deviation	N
VOCABULARY PRODUCTIVE TEST 1	Word	12.67	.617	15
	Sentence	12.73	.458	15
VOCABULARY PRODUCTIVE TEST 2	Word	12.13	0.990	15
	Sentence	12.27	0.961	15

As for the language of glosses, Figure 4.13 shows that the learners who had looked-up the English glosses did better in both tests 1 and 2.

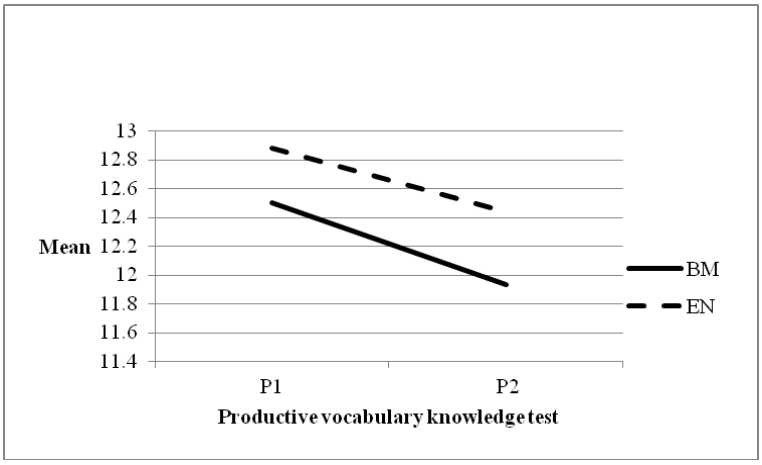


Figure 4.13: Language of glosses for productive knowledge test scores in high proficiency level

From Figure 4.14, it can be seen that the learners in this group scored better in the means of the test 1 and test 2 when using the sentence type of glosses.

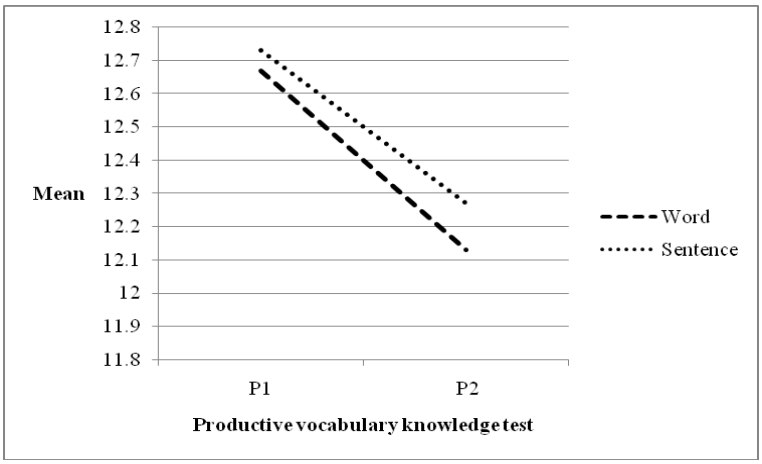


Figure 4.14: Type of glosses for productive knowledge test scores in high proficiency level

(a)Test of within-subject effects

The main effect for the within-subject variable of time was not significant, F (1, 26) =7.810, P=0.010, η^2 = 0.231 indicating that there were no statistically significant differences among mean tests score between test 1 and test 2. The results of Mixed ANOVA are summarized in Table 4.23**.** As can be seen in Table 4.23 all two-way and three way interactions (time*type, time*language and time*type*language) were insignificant.

Table 4.23: Results of Mixed ANOVA in high proficiency level

Source	df	MSE	F	p-value	η^2
Time	1	3.800	7.810	.010	.231
Time * Language	1	.067	.138	.714	.005
Time * type	1	.015	.030	.864	.001
Time * Language*type	1	.015	.030	.864	.001
Error (within subjects)	26	.487			
Language	1	2.917	3.938	.058	.132
type	1	.157	.213	.649	.008
Language*type	1	.024	.033	.858	.001
Error (between)	26	.741			

(b)Test of between-subject effects

The results of tests for between-subject effects indicated that for subjects in high proficiency level neither the language of gloss F (1, 26) = 3.938, P=0.058, η2 = 0.132 nor the type of gloss F (1, 26) = 0.213, P=0.649, η^2 = 0.008 had a significant effect on vocabulary acquisition. The language*type interaction was not significant.

This result showed that there were no significant differences in test scores between the two languages (BM and EN) and two types of gloss (word and sentence).

4.3.3 Receptive vocabulary knowledge

Again a mixed between-within ANOVA with two between-subject (language and type of gloss) and one within-subject variable (word receptive tests 1 and 2) was conducted to investigate the effect of language and type of gloss on vocabulary knowledge through word receptive tests.

This study incorporated the between-subject variables of language with two levels (BM and EN), the between-subject variable of type of gloss with two levels (word and sentence) and, the within-subject variable of time (word receptive test 1 and

word receptive test 2). The Mixed ANOVA tested whether there were the main effects for each of the independent variables and whether the interactions between the variables were significant.

Preliminary assumption testing was conducted to check for normality and homogeneity of variance of the mean test scores. Initial results showed that the two assumptions were fulfilled. Since the within-subject variable of time has only two levels (test 1and test 2) so it was not necessary to check the assumption of Sphericity for all proficiency levels.

4.3.3.1 Low proficiency

Tables 4.24 and 4.25 present the descriptive statistics for test scores for different gloss conditions.

Table 4.24: Mean and standard deviation for each language in low proficiency level

	Language	Mean	Std. Deviation	N
WORD RECEPTIVE TEST1	BM	8.077	1.605	13
	EN	7.294	2.845	17
WORD RECEPTIVE TEST2	BM	6.614	1.660	13
	EN	6.059	2.384	17

Table 4.25: Mean and standard deviation for each type of gloss in low proficiency level

	Type of gloss	Mean	Std. Deviation	N
WORD RECEPTIVE TEST1	Word	8.083	2.193	12
	Sentence	7.333	2.520	18
WORD RECEPTIVE TEST2	Word	6.500	1.567	12
	Sentence	6.166	2.407	18

Figure 4.15 shows that the learners obtained higher scores with interaction with the BM glosses than the English glosses across both test 1 and test 2, although this difference was not significant. In both gloss languages (BM/EN), the test scores decreased from test 1 to test 2.

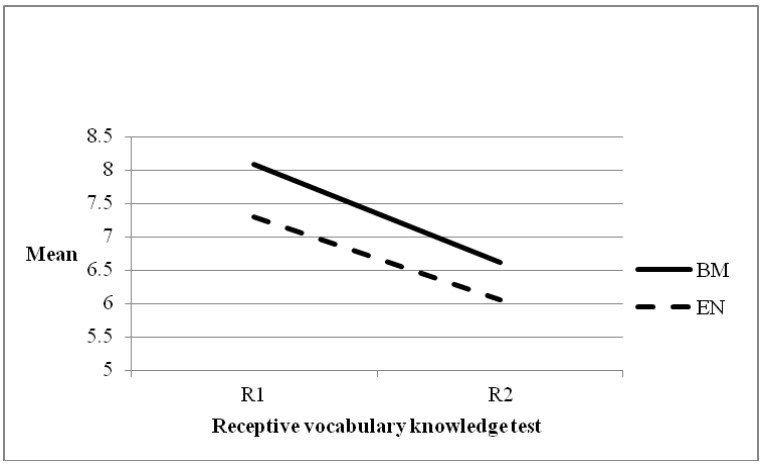


Figure 4.15: Language of glosses for receptive knowledge test scores in low proficiency level

Figure 4.16 shows that the learners obtained higher scores with interactions with the word glosses than sentence glosses across both test 1 and test 2, although this difference was not statistically significant (because the main effect of type was not significant). In both gloss types (word /sentence), the test scores decreased from test 1 to test 2.

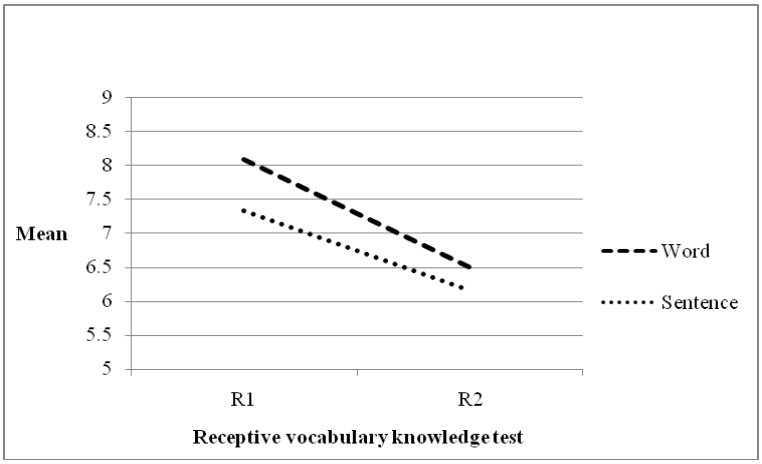


Figure 4.16: Type of glosses for receptive knowledge test scores in low proficiency level

(a)Test of within-subject effects

The main effect for the within-subject variable of time was significant F (1, 26) =10.995, P=.003, η^2 = 0.297 indicating that the mean scores significantly decreased from test 1 (M=7.63, SD=2.385) to test 2 (M=6.30, SD =2.869). The results of Mixed ANOVA are presented in Table 4.26. As can be seen in Table 4.26 all two-way and three way interactions (time*type, time*language and time*type*language) were insignificant.

Table 4.26: Results of Mixed ANOVA in low proficiency level

Source	df	MSE	F	p-value	η^2
Time	1	27.602	10.995	.003**	.297
Time * Language	1	.060	.024	.878	.001
Time * type	1	.480	.191	.665	.007
Time * Language*type	1	.312	.124	.727	.005
Error (within subjects)	26	2.511			
Language	1	5.303	.641	.431	.024
type	1	3.119	.377	.545	.014
Language*type	1	.012	.002	.969	.000
Error (between)	26	8.274			

(b)Test of between-subject effects

The results of tests for between-subject effects indicated that for subjects in low proficiency level neither the language of gloss F (1, 26) =0.641, P=0.431 η^2 = 0.024 nor the type of gloss F (1, 26) = 0.377, P=0.545, η^2 = 0.014 had a significant effect on vocabulary development. The language*type interaction was not significant. There was no significant difference in the test scores for the two languages (BM and EN) and two types of gloss (word and sentence).

4.3.3.2 Mid proficiency

Tables 4.27 and 4.28 present the descriptive statistics for test scores for different gloss conditions.

Table 4.27: Mean and standard deviation for each language in mid proficiency level

	Language	Mean	Std. Deviation	N
WORD RECEPTIVE TEST1	BM	10.4444	2.03563	18
	EN	9.3333	2.43584	21
WORD RECEPTIVE TEST2	BM	8.6111	2.22655	18
	EN	8.7619	2.21144	21

Table 4.28: Mean and standard deviation for each type of gloss in mid proficiency level

	Type of gloss	Mean	Std. Deviation	N
WORD RECEPTIVE TEST1	Word	10.4737	2.29416	19
	Sentence	9.2500	2.19749	20
WORD RECEPTIVE TEST2	Word	9.0000	2.40370	19
	Sentence	8.4000	1.98415	20

Figure 4.17 shows that the learners obtained higher scores under the BM gloss condition than under the English gloss condition across both test 1 and test 2, although these differences were not significant. In both gloss languages (BM/EN), the test scores decreased from test 1 to test 2.

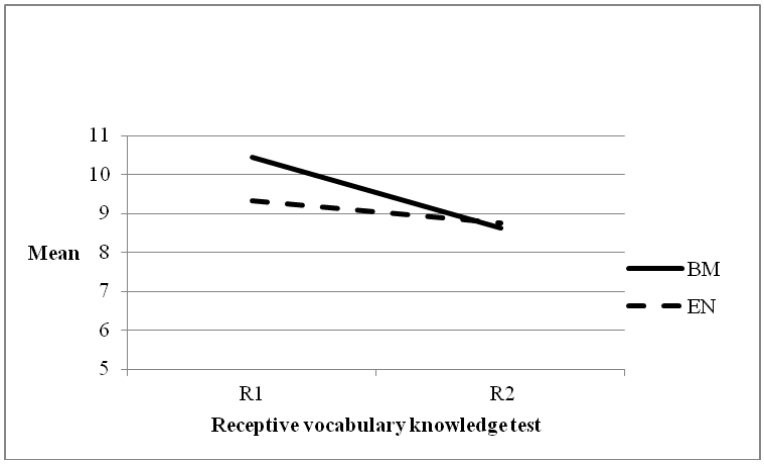


Figure 4.17: Language of glosses for receptive knowledge test scores in mid proficiency level

As can be seen in Figure 4.18 the learners who had used the the word glosses did better than learners who had used the sentence glosses across both test 1 and test 2, although this difference was not statistically significant (because the main effect of type is not significant. In both gloss types (word/ sentence), the test scores decreased from test 1 to test 2.

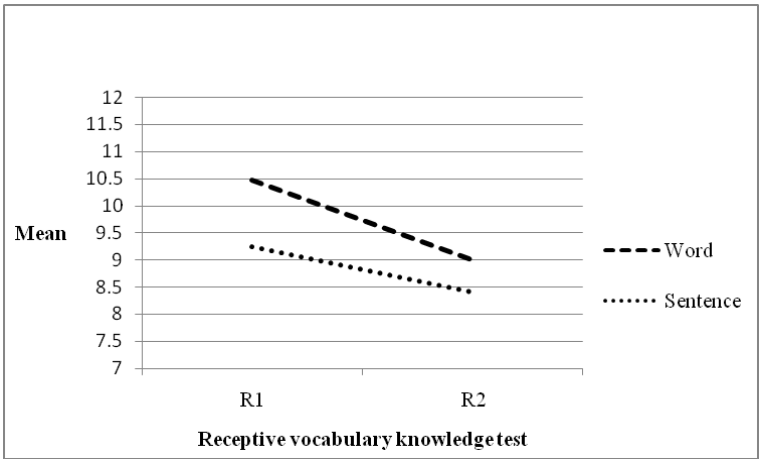


Figure 4.18: Type of glosss for receptive knowledge test scores in mid proficiency level

(a)Test of within-subject effects

The main effect for the within-subject variable of time was significant, F (1, 35) =20,719 P=0.000, η^2 = 0.259 indicating that the mean scores significantly decreased from test 1(M=9.85, SD=2.30) to test 2 (M=8.70, SD=2.19). As can be seen in Table 4.29 all two-way and three way interactions (time*type, time*language and time*type*language) were insignificant. The results of Mixed ANOVA are presented in Table 4.29.

Table 4.29: Results of Mixed ANOVA in mid proficiency level

Source	df	MSE	F	p-value	η^2
Time	1	41.211	20.719	.000**	.372
Time * Language	1	1.343	.675	.417	.019
Time * type	1	.145	.073	.789	.002
Time * Language*type	1	.708	.356	.555	.010
Error (within subjects)	35	1.989			
Language	1	10.594	2.111	.155	.057
type	1	24.906	4.964	.032	.124
Language*type	1	7.786	1.552	.221	.042
Error (between)	35	5.017			

(b)Test of between-subject effects

The results of tests for between-subject effects indicated that for subjects in mid proficiency level neither the language of gloss F (1, 35) = 2.111, P=0.155 η^2 = 0.057 nor the type of gloss F (1, 35) = 4.964, P=0.032, η^2 = 0.124 had a significant effect on vocabulary acquisition. The language*type interaction was not significant. Since the main effects of type and language and their related interaction effects were not significant, therefore there was no significant difference in the mean test scores between the two types of gloss (word and sentence) and the two languages (BM and EN).

4.3.3.3 High Proficiency

Tables 4.30 and 4.31 present the descriptive statistics for test scores for different gloss conditions.

Table 4.30: Mean and standard deviation for each language of gloss in high proficiency level

	Language	Mean	Std. Deviation	N
WORD RECEPTIVE TEST1	BM	11.9286	1.54244	14
	EN	11.5625	1.93111	16
WORD RECEPTIVE TEST2	BM	11.5000	1.45444	14
	EN	10.8750	2.09364	16

Table 4.31: Mean and standard deviation for each type of gloss in high proficiency level

	Type of gloss	Mean	Std. Deviation	N
WORD RECEPTIVE TEST1	Word	12.3333	1.17514	15
	Sentence	11.1333	2.03072	15
WORD RECEPTIVE TEST2	Word	11.6667	1.71825	15
	Sentence	10.6667	1.83874	15

Figure 4.19 shows that the learners obtained higher scores under the BM gloss condition than under the EN gloss condition across both test 1 and test 2, although this difference was not significant. For both languages (BM/EN), the test scores decreased from test 1 to test 2.

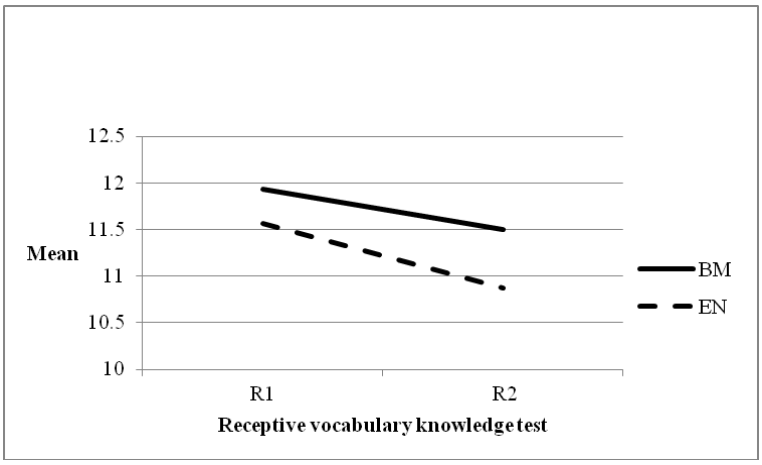


Figure 4.19: Language of glosses for receptive knowledge test scores in high proficiency level

In Figure 4.20, the learners obtained higher scores under the word gloss condition than under the sentence gloss condition across the both test 1 and test 2, although this difference was not statistically significant. For both gloss types (word/sentence), the test scores decreased from test 1 to test 2.

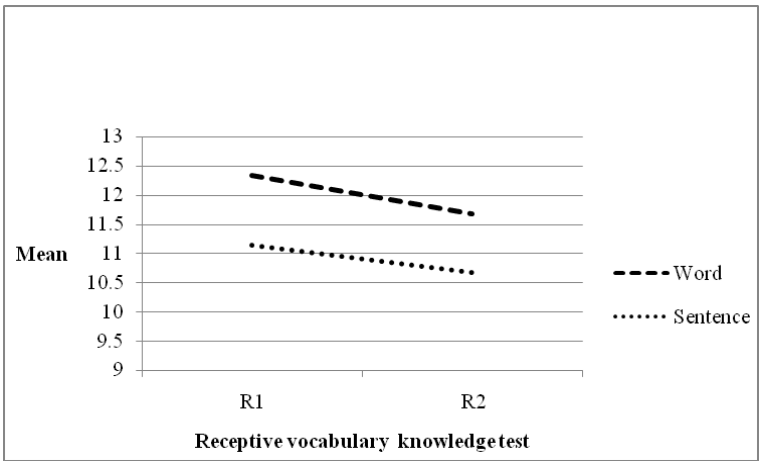


Figure 4.20: Type of glosses for receptive knowledge test scores in high proficiency level

(a)Test of within-subject effects

The main effect for the within-subject variable of time was not significant F (1, 26) = 2.559, P=0.122, η^2 = 0.090 indicating that there was no significant difference in test 1 mean scores and test 2 mean scores. The results of Mixed ANOVA are presented in Table 4.32.

Table 4.32: Results of Mixed ANOVA in high proficiency level

Source	df	MSE	F	p-value	η^2
Time	1	3.667	2.559	.122	.090
Time * Language	1	.067	.047	.831	.002
Time * type	1	.407	.284	.598	.011
Time * Language*type	1	.007	.005	.943	.000
Error (within subjects)	26	1.433			
Language	1	.967	.389	.538	.015
type	1	1.181	.475	.497	.018
Language*type	1	1.181	.475	.497	.018
Error (between)	26	2.487			

(b)Test of between-subject effects

The results of tests for between-subject effects indicated that for subjects in high proficiency level neither the language of gloss F (1, 26) = 0.389, P=0.538, η^2 = 0.015 nor the type of gloss F (1, 26) = 0.475, P=0.497, η^2 = 0.018 had a significant effect on vocabulary acquisition. The language*type interaction was not significant. As can be seen in Table 4.32 all two-way and three way interactions (time*type, time*language and time*type*language) were insignificant. To aid interpretation of the interaction effect, it would be useful to examine the graphs presented as Figures 4.19 and 4.20.

To summarise the impact of type and language of glosses on vocabulary knowledge by the students in different levels of proficiency, the results are tabulated in Table 4.33. From the table, it can seen that both the low and mid level proficiency students had similar gloss interactions. Initially for both these groups, BM and word type of glosses had aided them in their perceived vocabulary knowledge. After this, although BM glosses were still useful, sentence type of glosses were deemed more helpful for these students for perceived and productive type of vocabulary knowledge.

For high proficieny level students, on the other hand, showed a different pattern of interactions with the glosses. For initial perceived vocabulary knowledge, it was English and word type of glosses, moving on to English and sentence type of glosses for delayed perceived and productive vocabulary knowledge. In contrast, for all three proficiency levels, the BM and word type of glosses benefitted them in their receptive vocabulary knowledge.

Table 4.33: Type and language of glosses used by students of different proficiency levels and vocabulary knowledge

Proficiency	**Gloss**	**Imm PVK**	**Delayed PVK**	**Imm Prod. VK**	**Delayed Prod. VK**	**Imm Recep. VK**	**Delayed Recep. VK**
Low	Language	BM	BM	BM	BM	BM	BM
	Type	W	S	S	S	W	W
Mid	Language	BM	BM	BM	BM	BM	BM
	Type	W	S	S	S	W	W
High	Language	EN	EN	EN	EN	BM	BM
	Type	W	S	S	S	W	W

Note
Imm -immediate
PVK- Perceived vocabulary knowledge
Prod.VK -Productive vocabulary knowledge
Recep.VK - Receptive vocabulary knowledge
BM - Bahasa Malaysia
EN- English
W- word
S- sentence

4.4 Long-term gains in vocabulary knowledge

Research question 2(b):

Is the knowledge maintained over time?

Data gathered from the research showed that perceived, production and receptive types of vocabulary knowledge were lost three weeks after interactions with the glosses.

What is noticed was that although word and sentence glosses in L1 and L2 had helped the learners develop vocabulary knowledge of the perceived, production as well as receptive type of knowledge, it was not maintained after three weeks.

4.4.1 Perceived vocabulary knowledge

For the type of glosses, the results revealed that word and sentence type of glosses helped the students gain in their perceived vocabulary knowledge, however sentence type of glosses helped more in sustaining this type of vocabulary knowledge for all proficiency level students.

As for the language of glosses, it seemed that BM or L1 was more beneficial for the low and mid proficiency learners to gain and maintain perceived vocabulary knowledge. For high proficiency learners, it is clear that English glosses have the upper edge in aiding the learners to develop perceived vocabulary knowledge as well as to maintain it. It can be noticed also that there appeared to be a shift towards English glosses in maintaining vocabulary knowledge of learners with high proficiency level in the language.

4.4.2 Productive knowledge

For the low proficiency level learners, the test within-subject effects were insignificant meaning that although there was a decline in the productive knowledge, it was not significant. This meant that the low proficiency learners who had interacted with the glosses could somehow maintain the productive type of knowledge. It meant that the interactions with the sentence type of glosses and glosses in BM were able to aid this group of learners develop and somewhat maintain productive type of

vocabulary knowledge. For these learners, data showed that neither the language nor the type of glosses had a significant effect on productive knowledge gained. However, it on closer observation learners obtained higher scores in BM gloss conditions. The attrition was greater for English glosses than BM ones (EN gloss -1.11 and BM gloss -1.00). As for type, the low proficiency learners obtained better scores in Productive test 1 and Productive test 2 using sentence glosses. The attrition rate was greater for sentence gloss than word gloss conditions (Sentence -1.11 and word -1.00).

For the mid proficiency learners, the test within-subject effects were significant. This showed that the difference in mean for production test 1 and 2 was large. This meant that even though the learners had interacted with the glosses, there was significant attrition of productive type of knowledge for this group of learners. As for the test between-subject, it showed that language nor type of gloss had a significant impact on vocabulary knowledge. However, closer observation of the data showed that the BM glosses had a more positive impact on the learners' productive knowledge in both the tests. But, the attrition was higher for BM than EN glosses (BM -1.77 and EN -1.57). As for type of glosses, sentence glosses were more beneficial than word glosses and the decrease was also greater for word glosses than sentence glosses (sentence -1.6 and word gloss -1.73).

In the case of high proficiency learners, the test of within-subject revealed that the variable of time was not significant. This meant that the loss of productive knowledge was marginal. The results of the tests for between-subject showed that neither the language nor the type of the glosses had a significant effect on productive knowledge.

As for type of glosses, learners who had used the sentence type of glosses had better scores for both tests 1 and 2. The word gloss type had also a higher attrition rate compared to the sentence glosses (word -0.54 and sentence -0.46).

This group of learners displayed that for language of the glosses, it appeared that L2 or English had aided this group of learners. The decrease was greater for BM glosses than English glosses (BM glosses -0.57 and EN glosses -0.44).

4.4.3 Receptive knowledge

The test results of the test of within-subject showed that the variable of time was significant for low proficiency level learners. This indicated that the mean scores of W1 and W2 were different, with test 1 scores higher than test 2. This translated into a significant attrition of receptive knowledge for this low proficiency group of learners.

The test of between-subject showed that both variables language and type had an in significant effect on receptive vocabulary knowledge. Closer scrutiny however showed that for the type of glosses, it was apparent that learners who interacted with the word glosses scored higher means in the receptive test than sentence gloss conditions. Accordingly then, the sentence glosses loss was greater than for word glosses (sentence glosses -1.16 and word glosses -1.58). As for language of the glosses the learners scored better in BM gloss conditions rather than English glosses. However, the attrition rate for BM gloss conditions was also greater than the English glosses (BM glosses -1.46 and EN glosses -1.25).

For the mid proficiency learners, the test of within-subject revealed that the variable of time was significant. This meant that there was significant attrition of receptive knowledge for mid proficiency learners. The test of between-subject effects showed that the language or the type of gloss had no effect on the mid proficiency students. However, more concentrated analysis showed that for language of glosses, learners benefited more from interactions with the BM glosses than EN ones. Nevertheless, the loss was greater for BM glosses than EN glosses (BM glosses -1.83

and EN glosses -0.57). For the type of glosses, the results showed that learners had better gains after interactions with the word glosses compared to sentence gloss conditions. The decrease is greater for word glosses than sentence glosses (word glosses -1.47 and sentence glosses -0.85).

For high proficiency learners, the test of within-subject effects showed that the variable of time was not significant for these learners. This meant that the difference in test 1 and test 2 scores was not significant. This meant that learners in this proficiency level maintained the receptive type of knowledge over the three weeks. The test of between-subject effects displayed that the effect of language and type of glosses was not significant. Closer analysis would reveal that for type of glosses, it appeared that learners had gained more in word gloss conditions. The decrease was marginally more for word glosses than sentence glosses (word glosses -0.67 and sentence glosses -0.46). As for the type of language, the learners did better in the BM glosses. Moreover, the decrease was slightly higher in the English glosses than in the BM glosses (BM glosses -0.42 and EN -0.68).

4.5 Language proficiency, gloss use and vocabulary knowledge

Research question 3:

Is language proficiency of the learners a factor in determining how the learners interact with the glosses and subsequent effect on vocabulary knowledge?

In order to answer the first part of this research question pertaining to proficiency level and interaction, it is necessary to examine the data of research questions 1 to 2. Firstly, if the number of clicks with the target words were analysed, it was found out that the manner in which the different proficiency levels interacted with

the gloss was almost the same. However, a narrower observation of the means of the number of clicks revealed that the mid proficiency learners had the most number of clicks. The high proficiency learners had the least. This translated into that the learners with mid proficiency level interacted with the glosses the most. Secondly, as for the interaction with the type and language of glosses, both type of glosses (word and sentence) and language of glosses (BM and EN) aided learners in short-term vocabulary knowledge development. Closer examination however revealed that low and mid proficiency level learners benefitted from interactions with sentence type glosses in L1. On the other hand, sentence type of glosses in L2 made a positive impact on high proficiency learners.

The next part of the research question number 3 entailed an examination of the mean test scores of the students in the different proficiency levels and the type of vocabulary knowledge which is described next.

4.5.1 Perceived vocabulary knowledge

The mean sores and standard deviations for tests are provided in Table 4.34. It would be clearer to refer to Figure 4.21 to explain the data in Table 4.34.

Table 4.34: Mean scores and standard deviation for perceived vocabulary knowledge tests

Level of Proficiency		Pre-test score	Imm. test score	Delayed test score
Low	Mean	1.8889	5.6854	4.9907
	Std. Deviation	1.53478	3.10732	2.93259
	N	30	30	30
Mid	Mean	3.6364	7.6970	7.3408
	Std. Deviation	1.77966	3.20747	2.84081
	N	39	39	39
High	Mean	5.4595	9.6036	9.1895
	Std. Deviation	2.50967	2.64206	2.60140
	N	30	30	30

Figure 4.21 shows the interaction plot of the estimated marginal mean scores across pre-test, post test and delayed post test in different proficiency levels. This figure suggests that the high proficiency level group gained more than the two others group. But all three groups experienced the same trend (parallel lines), where there was increase in the immediate test followed by a decrease in the delayed test.

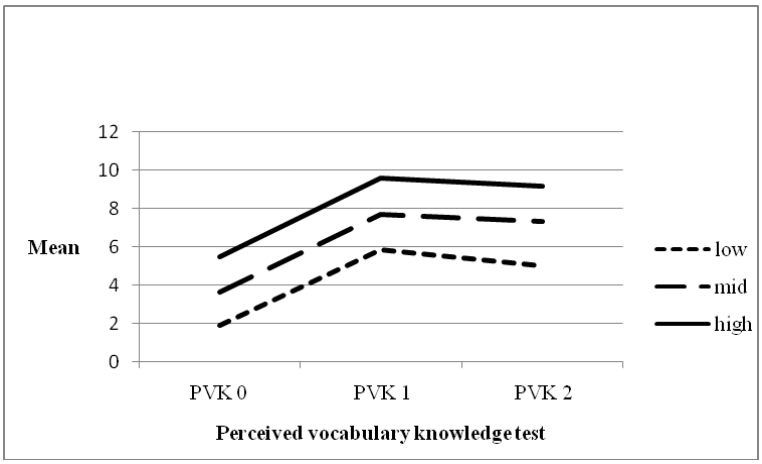


Figure 4.21: Mean scores of perceived vocabulary knowledge in different proficiency levels

4.5.2 Word productive knowledge

The mean sores and standard deviations for tests in each proficiency level are provided in Table 4.35.

Table 4.35: Mean scores and standard deviation for productive tests

	Level of Proficiency	Mean	Std. Deviation	N
VOCABULARY PRODUCTIVE TEST 1	Low	11.67	1.093	30
	Mid	12.21	.923	39
	High	12.70	.535	30
	Total	12.19	.965	99
VOCABULARY PRODUCTIVE TEST 2	Low	10.60	2.660	30
	Mid	10.54	1.958	39
	High	12.20	.961	30
	Total	11.06	2.104	99

Figure 4.22 shows the interaction plot of the estimated marginal mean scores across test 1 and test 2 in different proficiency levels. An inspection of the results and Figure 4.22 reveals that averaged across the three tests, those with high proficiency levels had the highest test scores , followed by those with mid proficiency level. The least test scores group was those with low proficiency levels.

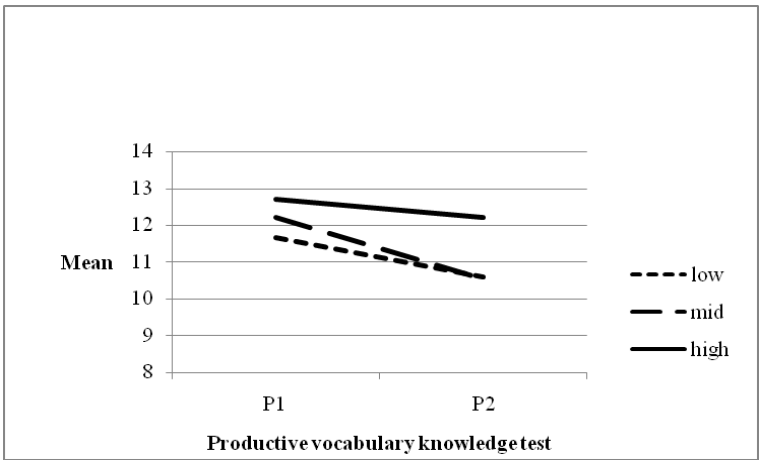


Figure 4.22: Productive test mean scores and proficiency levels

4.5.3 Word receptive knowledge

The mean sores and standard deviations for tests in each proficiency level are provided in Table 4.36.

Table 4.36: Mean sores and standard deviation for word receptive tests in each proficiency level

	Level of proficiency	Mean	Std. Deviation	N
WORD RECEPTIVE TEST1	Low	7.6333	2.38506	30
	Mid	10.1026	1.83238	39
	High	12.1000	1.29588	30
	Total	9.9596	2.55917	99
WORD RECEPTIVE TEST2	Low	6.3000	2.08690	30
	Mid	8.6667	2.09427	39
	High	11.6000	1.40443	30
	Total	8.8384	2.80927	99

Figure 4.23 shows the interaction plot of the estimated marginal mean scores across word receptive test 1 and word receptive test 2 in different proficiency levels. The results indicated that learners with high proficiency levels had the best test scores than mid and low levels, averaged across the two tests.

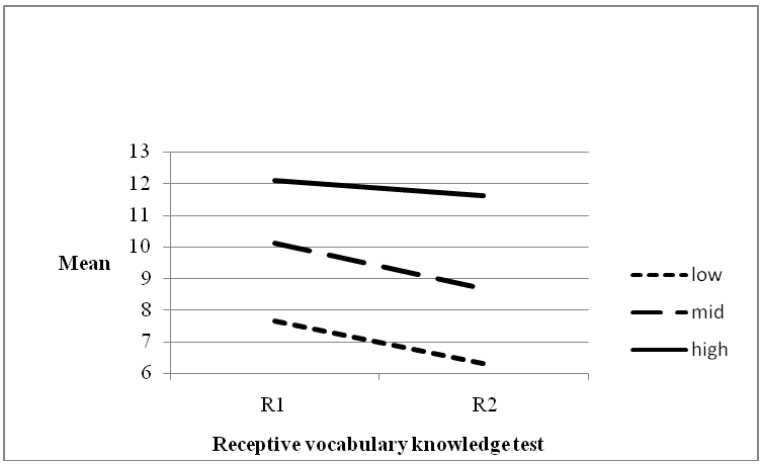


Figure 4.23: Receptive test mean scores and proficiency levels

As for vocabulary development, it is best to examine the three types of vocabulary knowledge which were measured in the study. Firstly, the perceived vocabulary knowledge, similar patterns of vocabulary development can be seen across all three proficiency levels, however, high proficiency learners made the most gains in vocabulary knowledge development in interaction with all gloss conditions.

The gains were significant across the two tests, that are pre, and immediate, and pre and delayed test, while no significant gains were recorded from the immediate to the delayed test. This means that glosses were effective in perceived vocabulary knowledge development

With the productive and receptive vocabulary knowledge development, the high proficiency learners benefitted the most from glosses in both types of vocabulary knowledge that is productive and receptive. Again what can be observed is that all proficiency levels had a similar pattern of development and type of vocabulary knowledge gained meaning that high proficiency learners made the most gains,

followed by mid and low proficiency level learners. However, the mid proficiency learners displayed a greater loss of productive knowledge compared to the low and high proficiency learners. Parallel to the patterns in productive vocabulary knowledge, the high proficiency learners benefitted the most from interaction with the glosses in receptive knowledge followed by the mid and low proficiency learners.

All three groups recorded a drop in the production and receptive delayed tests. For the low proficiency students, the drop was significant for receptive vocabulary knowledge, the mid proficiency level students experienced significant losses in both receptive and productive vocabulary knowledge. The high proficiency students had no significant attrition in both types of productive and receptive vocabulary knowledge. Therefore, this probably means that the high proficiency learners had gained and maintained vocabulary knowledge.

The mid proficiency level students had marked losses in both productive and receptive vocabulary knowledge, while low proficiency level students had significant losses in receptive vocabulary knowledge over time. Therefore it can be assumed that although there was vocabulary knowledge gained for low and mid proficiency learners, it was not sustained over time for these students in certain types of vocabulary knowledge.

It is also useful to examine the mean of the productive and receptive vocabulary test scores as tabulated in Table 4.37. From the mean of all three proficiency levels, it can be seen that mean of the productive vocabulary tests were higher than the mean of the receptive vocabulary tests. It is evident that students performed better in both the immediate and delayed productive vocabulary knowledge tests (P1 and P2) compared to the receptive vocabulary knowledge tests (R1 and R2). This meant that interaction with textual glosses may be able to aid learners in developing more productive type of vocabulary knowledge than receptive type.

Table 4.37: Proficiency levels and mean of productive (P) and receptive (R) vocabulary test

Proficiency level	Mean (P1)	Mean (P2)	Mean (R1)	Mean (R2)
Low	11.67	10.60	7.63	6.30
Mid	12.21	10.54	10.10	8.66
High	12.70	12.20	12.10	11.60

4.6. Conclusion

This chapter presented and analysed the data from the experiment. The data was analysed from the aspect of clicking and the specific interactions with the glosses by the students in the different proficiency levels. The interaction data were also examined in the light of the types of vocabulary knowledge gained or lost by the students. It was seen that clicking behaviour did not differ much from the students with different language proficiency levels. The type and language of glosses did not have a significant impact on the students' vocabulary knowledge development, although some patterns of interaction did occur when the data were observed closely. The data also showed evidence that textual vocabulary glosses in both *Bahasa Melayu* and English could not sustain vocabulary knowledge gained over time. As for proficiency levels, a predictable finding was revealed in which high proficiency level students gained the most from the glosses, followed by mid and low proficiency levels students. It was also found out that students gained more productive than receptive vocabulary knowledge. The ramifications of the learner-computer interactions on the Input-Interaction model and vocabulary knowledge development are discussed next in the following chapter.

CHAPTER 5

DISCUSSION AND CONCLUSION

5.1 Introduction

This final chapter of the thesis is divided into several sections. It begins with an overview of the findings of the study in relation to the existing literature in the area of gloss use, followed by implications for the mediating hypotheses and theory that underpin the study. It also deals with research and pedagogical applications of the study. Lasty, limitations of the study and recommendations for future research preceed the conclusions for this study.

Framed within the Interactionist approach and specifically utilising the Input-Interaction model of SLA in a CALL environment, this quasi-experimental study randomly placed learners with low, mid and high proficiency levels in different gloss conditions in order to investigate if the type of gloss, word or sentence; and the language of the gloss, L1 or L2 affected the learners 'vocabulary knowledge . In other words, the study investigated what kind of learner-computer interactions with modified input that could benefit the learners in their vocabulary knowledge. The research design was of a pre-test, post-test one and it was trajectory in nature as it tracked the development of the learners' vocabulary knowledge of the target words before the treatment (gloss use), immediately after the treatment, and three weeks after the treatment. From this sequence, the type of vocabulary knowledge whether it was perceived, receptive or productive type of knowledge, and whether it was gained or otherwise was also investigated.

5.2 Overview of main findings

This section discusses the main findings of the research and they are grouped according to the areas corresponding to the study's research questions: clicking behaviour, interactions with glosses, maintaining vocabulary knowledge, and proficiency level of students and gloss use.

5.2.1 Clicking behaviour

In human-computer interactions, a ubiquitous feature that most web-users are accustomed to when interacting with a computer is clicking. A major factor in any online interface is the act of clicking of the mouse and in computer interaction literature this is commonly referred to as clicking behaviour. The act of clicking usually leads to further engagement with the icons, images or the links in the interface.

The data on the clicking behaviour of the learners in the study suggest that clicking behaviour did not vary much between the different gloss conditions and proficiency levels. From closer inspection of the data on gloss type, it was observed that learners in the word BM gloss condition had clicked most on the target words, followed by the learners in the sentence BM gloss condition. The lowest number of clicks was made by learners in the word English gloss condition.

From the aspect of proficiency level, it was seen that high proficiency learners clicked the least on the target words, while mid proficiency learners clicked the most. The low proficiency learners were moderate in their clicking behaviour.

What has transpired is different from the expectation of gloss use where students who had clicked more would have had more gains from students who had clicked less. Additionally, it would have been expected that learners of low proficiency level would

click more on the target words to help them with the meaning of unfamiliar words to them. All these did not materialise in this study. Furthermore, perhaps the students' proficiency played a role in the clicking behaviour. High proficiency students who are good in the language probably clicked less as they had the advantage of learning words with less effort compared to lower proficiency level students.

In the literature of computer clicking, Murphy (1999) categorises web-users as "surfers" and "searchers". He defines the "surfer" as one who clicks from link to link mainly for amusement and it can be inferred that a "surfer" is one who clicks without any direction or purpose. On the other hand, there is the "searcher" who is focussed in seeking specific information. What this means is that a "searcher's" clicking behaviour is purposeful. The researcher is cautious of using these terms to label the students in this study as surfers and searchers. Certainly more specific inquiry into how the web user interacts with a web site is needed to reveal these types of clicking behaviour. This is also beyond the scope of this thesis. Nonetheless, both these traits were possibly manifested in the study where mid proficiency students had the tendency for "surfer" clicking behaviour, while high proficiency students displayed more purposeful clicking behaviour. This resulted in the mid proficiency students having more loss of vocabulary knowledge compared to low and high proficiency students.

What can be seen from the clicking behaviour of the learners in this study seems similar to clicking behaviour in other studies. For instance, Hulstijn, Hollander and Greidanus (1996) observed that advanced L2 learners when reading for comprehension did not interrupt their reading process by looking-up unknown words. They presumed that the concern of such readers were for comprehension and did not use their time and mental effort to look-up the meaning of unknown words. Again the same phenomenon was seen in a study by Hulstijn, (1993) who found that learners with greater vocabulary knowledge generally looked-up fewer words than learners who had less vocabulary

knowledge. The same situation existed in this study. The learners with higher proficiency level clicked less on the target words as compared to the low and mid proficiency level learners. If this kind of logic rules, then it would make sense that low proficiency learners would have clicked on more words if compared to the higher levels.

In this study, however, an anomaly seems to have occurred where the mid proficiency learners had clicked the most. What could have occurred with this group of learners is what Robb (1999, p. 98) terms as "click happy behaviour". The mid proficiency students in this study may have displayed surfers' characteristics as described by Murphy, (1999). Therefore, the learners of mid proficiency level clicked on the words more compared to the other learners given the ease and convenience of looking-up the target words which could have lead to superficial, short-term gains and no long-term ones.

The words were clicked as the students were aware of them and the glosses attended to, but what was lacking was perhaps the processing of the meanings of the target words. Processing here may refer to the process of making the form-meaning link (Van Patten, 2012). What has probably transpired was that the targeted words were clicked by these students because the words were made visually salient. However, this group of students superficially processed these words which may have led to poor retention of vocabulary knowledge.

5.2.2 Interactions with glosses

For all proficiency levels, the results implied that there were no differences between the languages (BM & English) of glosses or between type (word & sentence) glosses in developing the different types of vocabulary knowledge.

A sharper analysis of the results however showed that BM and sentence type of glosses were more useful for low and mid proficiency students in maintaining perceived, and productive vocabulary knowledge while English and sentence type of glosses were useful for high proficiency students. As for receptive vocabulary knowledge, it was seen that there was a shift to word and BM glosses for students in all proficiency levels to gain in this type of vocabulary knowledge.

As for interaction with specific type and language of the gloss, it was found that both types of gloss, that is word and sentence in L1 and L2 were able to help learners develop vocabulary knowledge, be in perceived, productive as well as receptive in the short term. It was found that glosses at word and sentence levels, in both L1 and L2 can help promote vocabulary development immediately after interaction.

The discussion next moves to deliberate the type of interactions on vocabulary knowledge. To maintain perceived vocabulary knowledge, it seemed that sentence type of glosses in L1 would benefit learners with low and mid proficiency levels. There was a shift in the type of interaction with regards to high proficiency level learners. They appeared to have gained when interacting with sentence and glosses in L2 or English.

Interactions with sentence type of glosses and in L1 again aided learners in low and mid proficiency levels to make immediate gains in productive type of knowledge. At the next level of analysis (delayed), it was apparent that the same type of glosses can aid productive type of knowledge. As for high proficiency learners, sentence and L2 was observed to aid this group of students for both immediate and delayed productive vocabulary knowledge gain.

The interactions with textual glosses, in particular word type glosses apparently promoted receptive type of vocabulary knowledge in the short-term for all learners regardless of proficiency levels. There was a total shift in the interactional patterns when compared to productive type of vocabulary knowledge.

It emerged that for all proficiency levels, learners were helped by word type of glosses in BM or L1, even for the high proficiency students who previously had benefitted from sentence and L2 glosses.

5.2.2.1 Type of gloss

As for type of gloss, it appeared that all learners in low, mid and high proficiency groups benefitted from sentence type of glosses. This could mean that meanings of words which are contextualized in sentence forms aided learners in their vocabulary development. It could be said that with sentence type of glossing learners may have gained more as they had gone through a deeper processing of the sentences rather than just glancing at the word level type of glosses (Cheng & Good, 2009). It could also mean that meanings explained in the sentence form helps learners clarify the meaning more clearly from the context as it would happen in an oral setting or classroom environment where learners are able to get clarifications and/or confirmation checks of their linguistic shortcoming from the more able interlocutor or teacher.

The results from the study indicated that learners who had interacted with the sentence-type glosses had gained in vocabulary knowledge. This is consistent with other studies done on glossing (Chun & Plass, 1996; Chen and Good, 2009; Cumming, Cropp & Sussex, 1994; Grace, 1998; Hulstijn, 1993; Knight, 1994; Laufer & Hill, 2000; Lomicka, 1998) showed that sentence-level definitions of words are helpful for learners. Therefore, it can be said that interaction with sentence type of glosses is more beneficial for learners as it perhaps provided more context to the meaning of the words which enables the learners to learn vocabulary. However, this does not mean that word glosses do not aid the students as word type of glosses were beneficial for receptive vocabulary knowledge as seen in this study.

From the literature on textual glossing, it has been documented that word type of glosses do also have its merit. For example, in Davies and Lyman-Hager's (1997) study on glosses in French reading, subjects tended to utilize almost exclusively on word definitions provided in English (L1) ignoring other forms of glosses available. In their earlier study, Lyman-Hager and Davies (as cited in Laufer & Hill, 2000) suggested that accessing "word meanings in the native language is a key factor in comprehension other studies (Chun, 2001; Gettys, Imhof & Kautz, 2001 and Hegelheimer, 1998) have shown that word glosses were useful.

In this current study, it was observed that interactions with word-type of glosses had made the learners gain receptive vocabulary knowledge. What was more illuminating from this current study is that interactions with word type of glosses had aided the learners with different language proficiency levels to develop their receptive vocabulary knowledge. There seemed to be a corresponding match between the word type of gloss and receptive vocabulary knowledge. A probable reason for this would be the nature of the vocabulary tests, where the test format required learners to recognise word meanings devoid of any context. This matched the word type of glosses which only provided word or definitional meanings. This may mean that access to the word type of glosses provided interactions where the meanings of the target words are in a decontextualised manner, thus enabling learners to respond more effectively to receptive type of tests of the same nature.

It does appear then that if receptive vocabulary knowledge is tested in a vocabulary test of the similar nature, that is, providing the meaning of the word devoid of context such as in multiple choice receptive vocabulary test will show gains of receptive vocabulary knowledge. Therefore, it may mean that the type of interactions with the gloss type has to be directly linked to the format of the test and what type of knowledge gain is expected.

Although this may augur well for developing vocabulary knowledge, the possible drawback here is that the "Naïve Lexical Hypothesis" (Bland, Noblitt, Armington & Gay, 1990) is reinforced where learners assume that there is one-to-one correspondence between the words and their meaning. This is certainly not the case for vocabulary where words are used in different contexts and can have different meanings.

5.2.2.2 Language of gloss

In this study, in the interaction with the languages of the glosses, it was discovered that L1 type of glosses brought about immediate gains to learners in low and mid proficiency learners in their perceived, productive and receptive vocabulary development. Only the high proficiency learners had gains when using the L2 in perceived and productive types of knowledge. Overall, it emerged that interactions with glosses in L1 had benefitted the low and mid proficiency learners in developing their vocabulary knowledge.

In the literature of the language of glosses, it is apparent that there is variation on the benefits of using L1 or the L2. Jacobs et al., (1994) found that there was no significant difference between Spanish (L2) and English (L1). Laufer and Hill (2000) in their study showed that Israeli students' look-up behaviour slanted towards their L1, while the Chinese students preferred English (L2) look-ups.

In Ko's (2012) study, it was seen that English (L2) showed significant effects on students' comprehension but at the same time students indicated through a think-aloud protocol that Korean (L1) also helped facilitate comprehension. From Yoshii's (2006) study, it was discovered that both Japanese (L1) and English (L2) were effective for vocabulary learning, but long term retention may differ between the two languages.

Students who had used the L1 remembered the target words better than in the L2. Miyasako (2002) on the other hand revealed that English (L2) glosses were more effective than Japanese (L1) especially if the students have a higher proficiency level in the L2. From these findings, it is postulated that the L1 and L2 can be used to help students learn vocabulary.

Clearly, from the literature it was seen that the use of L1 and L2 in glosses can aid students to learn L2 vocabulary, and it seemed that the range of L1 may be different but the effects are still the same, that is it can be used for vocabulary learning purposes. This is in line with the findings from this study too where there was no significant differences in the use of BM or English glosses for vocabulary development. This could mean that both BM and English can be used for developing students' vocabulary knowledge. Furthermore, this study has probably contributed another L1 which is *Bahasa Melayu* or Malay to the range of first languages that can aid students in learning vocabulary. This may fill in the gap of languages in CALL as Felix (2008) reports that the prominent languages in the field have been Japanese, Chinese and Russian.

There is enough evidence to show that L1 and L2 type of glosses generally has benefits for the learners. Previous research has indicated that it does not matter whether the gloss is an L2 description of an L1 translation, as long as the learner can understand the meaning (Jacobs et al, 1994; Yoshii, 2006). However, there is a complication as there are two dissenting opinions on this issue. One is that there is little language development if L1 is used on the grounds that there is too little effort on the part of the learner to learn the meaning of the words, hence the words and their meanings are not processed deep enough for language development (Taylor, 2006).

On the upside, Taylor comments that when L2 readers understand more words with L1 glosses, they would be able to focus on other reading strategies such strategies such as activating background knowledge, analyzing text structure and using semantic-

mapping strategies. This would directly mean that their comprehension would be enhanced. In contrast, if learners are exposed to only L2 type of glosses, they will perhaps think harder about the meaning and therefore learn the words and probably this learning can be maintained over time (Stoehr, 1999). The implication of deep processing of vocabulary items may be at the expense of comprehension because the cognitive effort of the students will be placed on vocabulary learning.

This may mean that the choice of the language of the glosses would depend on the focus of the language development, that is, the focus on comprehension or vocabulary learning. What can be surmised from the earlier studies and this present one is that using glosses in the students' L1 and L2 have benefits for the students which may bring about vocabulary learning and perhaps enhanced comprehension.

5.2.3 Maintaining vocabulary knowledge

This section discusses the issue of maintaining vocabulary knowledge in relation to proficiency levels. As for sustaining vocabulary knowledge, there were no significant changes for the low proficiency students in both the perceived and productive type of vocabulary knowledge. However, it was discovered that there was significant drop in receptive type of knowledge for this group of students.

For the mid proficiency students, only perceived vocabulary had no significant changes. The students were able to maintain their perceived vocabulary knowledge with the use of glosses. However, it appeared that these students had significant losses in both productive and receptive vocabulary knowledge in the long term. It was discovered that the knowledge gained was not maintained over time (three weeks). What this suggests is that interactions (textual-only) of this nature may not benefit these students in the long term.

There appears to be a need for richer and more engaging type of interactions which induces more processing within the minds of these mid proficiency learners. This type of engagement may not be limited to mid proficiency students only. Processing would mean making the connection between the lexical forms and meaning (Van Patten, 2012).

Students with high level of proficiency had no significant changes in all three types of vocabulary knowledge. This is interpreted as this group of students managed to retain the vocabulary knowledge gained from the interactions with the textual glosses.

There are a few factors which can be put forward that may cause vocabulary knowledge to decline. One factor is there should be multiple exposures to the target words so that the knowledge of the meanings of the target words can be maintained. Exposures could be in the form of seeing the target words again in the context, or even in the tasks. Although the learners in this study may have clicked on the target words a few times, it may mean that exposure to the same format and type of glosses will not make a difference. There has to be multiple formats of exposure.

The other factor which needs to be utilised further in a computer environment is the provision of feedback from the computer to the learner. In this study, the feedback is the meaning of the target words embedded in the glosses provided positive feedback. This may not have been effective as it did not provide more processing by the students. In reflection, negative feedback to the learners could have resulted in them processing the meanings of the words more which could have brought about retention of the vocabulary knowledge gained. What this means is that the gloss could have been devised with various different meanings of the words to the learners. The learners then choose the best meaning which corresponds to their perceived meaning. If this is incorrect, the computer then provides alternative answers.

In this way, negative feedback may promote more processing on the part of the learner. Therefore, more modified interactions can take place between the learner and computer, thus offering more opportunities for negotiation and learning, possibly leading to retention. There have been advancements made in this area of enabling computer to provide both positive and negative feedback via the computer and this should be taken into account when designing CALL materials.

In relation to maintaining vocabulary knowledge what has been found out was similar to what has been reported in Watanabe's (1997) research where it showed that words which have been previously learnt would fade away if there was no reviewing process. Schmitt (2010) also made a similar claim where vocabulary attrition is a common occurrence in learning, more so for vocabulary. Hence, it did seem that this same situation may have been played out in this present study. Therefore, this researcher concurs with Pimsleur (1967) and Cheng and Good's (2009) view that new vocabulary items have to be reinforced by frequent reviews after they have been presented in order for the items to be retained over time.

In the case for low proficiency learners, the retention loss is consistent with the study by Abraham (2008) who also discovered that beginner L2 learners experienced a significant amount of vocabulary retention loss. The findings in this study are similar to the findings in Watanabe's (1997) study. In that study it was found out that even with glossing, learners with small vocabulary size would not be able to effectively use the glosses provided. It could be interpretated that low proficiency students were the ones who had small vocabulary size. What is apparent in this study is the mid proficiency learners had the highest attrition rate for both production and receptive learning among the three different proficiency levels. This is indeed peculiar as to why mid proficient learners seem to lose out the most in both productive and receptive learning in all gloss conditions.

A plausible explanation for the drop of the mid proficiency learners in the production knowledge is perhaps productive learning took more time than receptive learning (Nation, 2001). Hence the productive delayed test scores were lower than the immediate productive test scores. This means that the glosses that the learners had consulted did not offer the opportunity for more time to learn the targeted vocabulary. Therefore, the findings of the study resonates with Nation's assumption that "more time and repeated effort is needed to learn vocabulary for speaking and writing (productive) than is needed for listening and reading (receptive)" (p. 33). Although repetition is offered by the glosses, this facility may have not been utilized by the learners, particularly the mid proficiency level learners in the study, resulting in poor productive vocabulary scores.

What is important from this observation is best captured by Nation as, he explains that it is still not clear if readiness for productive use can be reached by receptive 'over-learning', or from Swain's (1985) point of view on whether there must be 'pushed' output with learners being made to speak or write. In other words, the glosses in the study did not provide the necessary 'push' nor the receptive learning needed for the learners to produce the language. What is needed perhaps tasks with such objectives of allowing the learners to produce the words which would lead to the words being acquired productively.

This mid proficiency learners also had high attrition rates for receptive type of knowledge. Although this group of students had clicked the most on the glosses, what had taken place was probably their "click-happy" and "surfer" behaviour may have surfaced which in turn led to low processing of the words, hence the sharp decline when compared to the low and high proficiency students. The researcher is aware that the terms linking such behaviour may have to be validated by more computer look-up behaviour research to warrant the use of the terms.

It appears that this type of clicking behaviour had given limited or superficial exposure of the glosses to this particular group resulting in them failing to grasp the meanings of the target words efficiently to maintain the knowledge across time. Still on this issue of maintaining vocabulary knowledge, it can be contended that the glosses may have not been processed in the working memory.

Ellis (2008) records the importance of working memory where "key processes of perception, attention and rehearsal take place" (p. 407). He also assumes that working memory is limited in its capacity. Therefore, for any kind of long term learning to occur, knowledge has to be stored in the working memory after the initial processing in the short term memory. In order to explain the link between processing and memory, the discussion turns to another model of the working memory provided by Baddeley (1986).

Baddeley proposed three different components which are a phonological loop and a visual/spatial sketchpad. He speculated that learners would gain if multiple resources are used to make sense of meaning and form through auditory and visual mediums. Hence a learner may gain in learning if he or she receives one aspect of the language in one medium while he or she learns another aspect of the language from another medium, for example, attending to form and meaning via different mediums. It can be assumed that multiple platforms should be made available to students for language learning. The consequence of such an interpretation is that if computer glosses are to be effective, they have to be designed using perhaps multimedia characteristics. To illustrate, to provide form and meaning using different channels such as audio, pictures, images or videos.

Given the centrality of the working memory and processing, it would seem that there has to be attempts to reinforce the processes in the working memory in order for it to be effective.

What may have occurred in this present study is that the textual aspect of the glosses did not provide enough push for the vocabulary to be held in the working memory. What is needed again is more processing by the learners by utilizing both the phonological loop and visual sketchpad in the form of visual and auditory glosses in order for the vocabulary items to be held longer in the working memory which may then be passed into long term memory corresponding to Gass' integration stage of the Input-Interaction model. This is what is meant by the interactions have to be multi-faceted for the learners to sustain vocabulary knowledge over time.

5.2.4. Proficiency level and gloss use

The analysis revealed that high proficiency learners gained the most in gloss use. They did not significantly lose vocabulary knowledge over time. Low proficiency learners had moderate gains; they lost receptive knowledge in the long term. In contrast mid proficiency learners had the most significant loss in gloss use, especially for productive and receptive type of knowledge over a period of three weeks.

As for types of interactions and proficiency levels, it looked like interactions with L1 glosses for low and mid proficiency levels supported immediate vocabulary knowledge gain. Look-ups with sentence type of glosses also helped all proficiency levels for vocabulary development, specifically for the perceived and productive vocabulary. For word receptive knowledge however, word type of glosses provided more support for learners in all the three proficiency levels. It can be summed up that learners with different proficiency levels benefited differently from the interactions with the glosses.

Another point on the interactions of the learners with different language proficiency levels was that the mid proficiency learners were unable to reap the benefits of the interactions with the gloss over time. They lost out on maintaining vocabulary knowledge, both production and receptive vocabulary knowledge. Despite being the group which clicked the most, their loss in vocabulary knowledge is peculiar. As discussed earlier, it could have been their "click-happy" and "surfer" behaviour. Also perhaps, this group of learners are at a developmental stage, where they were unable to gain from the use of glosses. In contrast, the low proficiency level learners benefitted from the interactions as they were able to maintain some form of productive knowledge while losing on receptive vocabulary knowledge in the long term. The low proficiency learners may have been at the right developmental stage for language learning and being prompted by the glosses could have resulted in them giving more attention to the glosses enabling some vocabulary knowledge gains.

The data also showed that the high proficiency level learners gained the most from the interactions with the textual glosses. Although there was attrition in the three types of vocabulary knowledge, it was not significant for the high proficiency learners. These learners were already good in the language. Hence, even with minimum number of clicks and interactions they were able to make gains.

From the language of glosses, it can be surmised that the low and mid proficiency level learners still needed the L1 as a crutch in the glosses. It looked like there was a language threshold before the learners could gain from the English glosses. Again with reference to studies by Ko (2005) and Taylor (2006), it was found out that learners would consult and use L2 glosses when the learners found them to comprehensible or the language used was pitched to their level. As for retention, it was seen that vocabulary knowledge was not maintained even if learners had used L1 glosses.

From the findings, it can be concluded that language proficiency is a factor in gloss effects. This is in line with the finding from Cheng and Good (2009). In their study, all proficiency levels gained from the use of glosses but the benefits were not equal. Here, it was seen that high proficiency learners gained the most, followed by mid and low proficiency learners. This observation does seem to concur with Watanabe's (1997) explanation that even with glosses which are comprehensible, learners with small vocabulary size would not benefit. Of course, in his study's context a small vocabulary size may be assumed to mean low proficiency. The higher proficiency learners performed better in the vocabulary tests could also be attributed to their competence in the L2 and they were able to maximize the use of the glosses for their benefit. This is similar to what Jacobs et al. (1994) had found in their study. The results from Jacobs' research are consonant with what has been found in this study where high proficiency students made the most gains in vocabulary knowledge compared to the mid and low proficiency students.

What can be learnt from the results is that it is important to consider learners' proficiency levels when designing glosses as learners in varying language proficiency levels benefit from interactions with glosses differently. Indeed, the researcher discovered that high proficiency level learners gained the most in perceived vocabulary knowledge as well as productive and receptive knowledge. This finding fits into the following observation by Groot (2000),"One might argue that high level learners have meta-cognitive strategies at their disposal which make their acquisition of new vocabulary much less dependent on externally imposed learning conditions than is the case for low level learners" (p. 21).

It is also documented (Cheng & Good, 2009; Jacobs et al., 1994) that high proficiency students with bigger vocabulary would consult lesser words compared to students in low proficiency with smaller vocabulary.

Although the high proficiency students made fewer consultations with the glosses, they were able to take advantage of the sentence type of glosses which provided more context for vocabulary development. This is in line with Schmitt's (2010) comment that language proficiency determines to what degree learners can take advantage of any contextualization in language learning tasks or tests. Clearly, in this study it has been demonstrated that high proficiency level students were able to benefit from the contextualization provided by the glosses.

Finally, it is difficult to extract the exact role of proficiency in relation to interactions. This is more apparent in gloss studies as the links between proficiency, type and language of glosses form an intricate web. Nonetheless, it is clear that proficiency, interactions with different glosses impact vocabulary knowledge development in different ways. The next section of the discussion deals with the implications of the study to the Input-Interaction model including the notions of attention, noticing and interaction.

5.3 Implications to theory: The Input-Interaction model

To recap, the study has looked at input in the form of an online text and the focus is on interaction of the students with computer textual glosses. This forms the theoretical framework in this research as it is drawn up within the Input-Interaction model by Gass (1997) which has been documented as a viable model in SLA. Ellis (2008) states that Gass' model "constitutes the fullest and clearest statement of the roles played by input and interaction in L2 acquisition currently available" (p. 268). This model is used to explain language acquisition by examining the notions of input and interaction. Ellis (2008) considers Gass' model of SLA as a "simple and serial-processing" one (p. 407).

It is simple in the sense that it looks at SLA as a process which begins with input which undergoes several phases of processes occurring at a time. These processes also occur in parallel, hence the term "serial-processing". It can also be interpreted as being computational in nature as it assumes that learners' language learning process is similar to that of a computer, in the sense that there are separate components which are related to the process of language learning.

In order to relate this model to the study it is important to consider the assumptions in the model. Eysenck (as cited in Ellis, 2008) essentially identifies several assumptions in this particular model. Firstly, there is input which is obtained from the environment. Secondly, there is cognitive processing where learners process the input in systematic ways. Thirdly, this processing involves progress into the working memory and long-term memory, and finally, the input processing that takes place in this model is similar to that in computers.

In relation to this study, from the online text input, cognitive processing took place which included the notions of attention, perception and short-term memory. At this stage the saliency of the target words in the text was operative. As the target words were highlighted in a different colour from the text, this created visual saliency which directed the learners to be aware and attend to the words by clicking on them. When clicked the words revealed the modifications that had been done on them. In total, the learners were able to obtain modified input in the form of glosses. The key to this type of glosses was that it was made interactive to the learners whereby learners can read the text and request for input modifications as they needed it. Chapelle (2003) explains that "the potential for such modifications is often seen in CALL materials which provide input that the learner can request it to be modified" (p. 58). This statement by Chapelle lends support to the study which examined the learner-computer interactions (with glosses) to facilitate vocabulary knowledge development.

That act of clicking set the motion of interaction with the glosses. With the interactions, learners would then have noticed the difference between what they knew of the meaning of the clicked words as compared to the meanings which were offered by the glosses.

This realisation corresponds to the apperceived stage of the model and it is an important one as the notion of noticing occurs in this stage. Schmidt (1990) claims that noticing something in the input is crucial to acquisition. This was also where the interactions with the glosses took place with the students obtaining modified input in the form of the glosses. This input was processed in the minds of the learners and the modified input is further processed into working memory and long-term memory, and the final assumption was that learners processed the input as a computer would – in a serial manner.

It was clear that the input which was apperceived and later processed had undergone mainly cognitive processes which were neither observable nor documented in this study. What was recorded was the clicks made by the learners which the study assumed signalled that the input had been apperceived. The subsequent process of the input was stored into the short-term memory which may had been made evident in the immediate vocabulary tests of the study. The results of the immediate tests were positive which suggested that the modified input had been comprehended, implying that the interactions with the glosses had a positive impact to the students. Clearly, there had been intake of the words by the students.

In this study, noticing was linked to interactions where the clicks made by the learners on the target words signalled attention and possibly subsequent noticing was given to words and their meanings. However, this aspect of attention and noticing were not probed further to find out if noticing was triggered by internal gaps in the language or just ordinary clicking behaviour triggered by external factors as the words in the text were highlighted.

This kind of probe would have required the use of protocols that could tap on the learners' mind as it processed the targeted words which the present study did not utilise.

As for modified input, it is observed from this study that textual glosses can help in developing vocabulary knowledge but not in maintaining it in certain cases. What is suggested is that textual glosses as modified input can help in the aspect of short-term vocabulary knowledge gains.

The output in the study was measured by the vocabulary tests which operated on the premise of the Comprehensible Output Hypothesis. Output as in vocabulary knowledge was measured immediately after the glosses were used and three weeks after the experiment.

The vocabulary results depicted that glosses were helpful in general to developing vocabulary knowledge but the help was limited as the knowledge gained was not durable after three weeks. The study also showed that vocabulary tests may not be an appropriate measure of vocabulary knowledge. This suggests that more meaningful tasks could better measure vocabulary knowledge instead of direct tests of vocabulary. What can be gleaned from the study was also the placement of the tests in the Input-Interaction model was too early to test output. The process of interaction should be more protracted and iterative before output is measured. This would probably provide a more reliable indicator to the whole process of language development as output is measured at the end of the processing continuum.

From the explanation given above, the implication for the Input-Interaction model is that the whole range of the Input-Interaction-Output perspective has to be in place. What this means is that the elements of attention, noticing, interactions with modified input and output has to be investigated in tandem to explain language development.

To bring this discussion on the Input-Interaction model to a close, it can be summarised that the notions of noticing and attention which preceded the apperception stage of the Input-Interaction model were manifested in the study. This led to the input being noticed or apperceived. The glosses which provided the modified input and interaction resulted in the students comprehending the input which led to intake. The immediate vocabulary results showed positive development of vocabulary knowledge. However, it could be seen from the results of the delayed vocabulary tests that this intake did not progress into full integration by the students. It appeared that there should have been more interactions that allow more processing within the learners' minds before effective integration can take place. Alternatively, more tasks could have been designed to reinforce the vocabulary learnt from intake into the integration stages.

The earlier figure of the Input-Interaction model (Figure 2.1) in Chapter 2 which shows the Input-Interaction model is redrawn as Figure 5.1 to show how this study infused the different components of the research into the model. The dotted lines denote where the components of this research fit into the Input-Interaction model of SLA. In Figure 5.1, it is shown that the input which is the online text had its target words modified and made salient. This visual saliency could have triggered the students to click on them. This act of clicking may have entailed the notions of awareness, attention and clicking. When clicked the students obtained modified input on demand. This triggered their interactions with the glosses which could have helped in making the input more comprehensible, in the sense that the meanings of the unfamiliar target words were made clearer to the learners. The immediate vocabulary tests measured the students' vocabulary knowledge immediately after the use of glosses possibly indicating the intake of students' vocabulary knowledge, while the delayed tests measured the students' vocabulary knowledge after three weeks; testing whether

the vocabulary knowledge of the students had been integrated and sustained. Finally, output was the overall vocabulary development of the students.

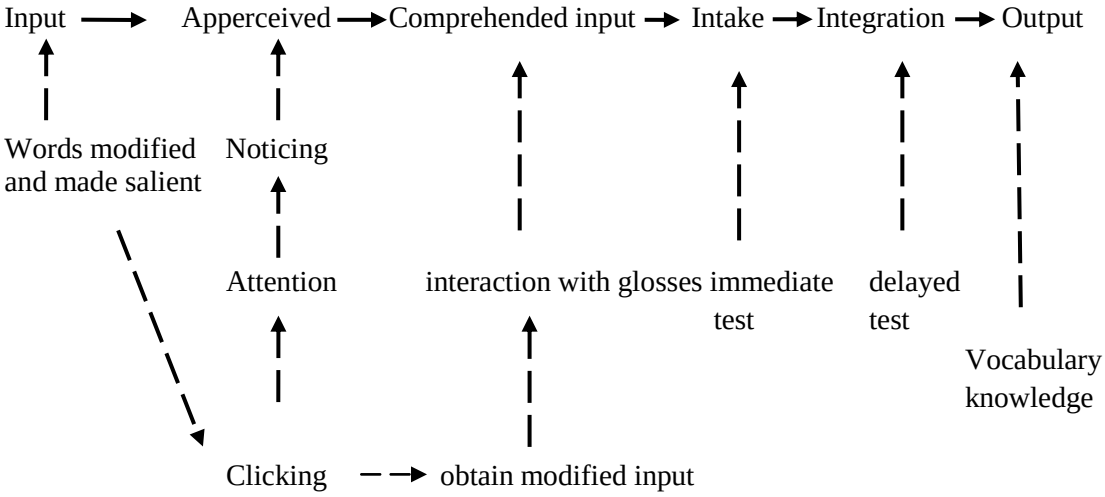


Figure 5.1: The Input-Interaction model as represented in the study

5.3.1 Attention

In SLA theory, conscious attention is an important role for language learning where learners notice the gap between what they say and the input they receive from the native speaker or the more able interlocutor (Fischer, 2007; Pica, 2005; Schmidt, 2001). It is this awareness of the mismatch that makes way for the language acquisition. Pica (2005) pointed out what is really important is not whether learners notice this mismatch but whether what they notice in input becomes intake, leading to changes in the interlanguage which is then internalized.

This study like other assessments in SLA used observed data as a basis for evidence of awareness and attention. In the study the clicks made by the students was used to make an inference on something which is not observable (Chapelle, 2003). In this present research, there was conscious attention when learners clicked on the target words.

After consulting the gloss, they may have noticed the gap in their own perceived meaning of the words as opposed to what has been provided by the gloss. This is what Pica wrote as "mismatch".

The question is then whether this difference becomes intake which is internalized is what that matters for acquisition. That part of the equation is certainly beyond this thesis as this process is largely within the learners' mind. Therefore, what can be shown here is that conscious attention is triggered by the words which have been made salient, prompting the learners to click on them, thus making a point for awareness which provides attention leading to noticing. This 'sequence' matches the stand taken by Gass and Selinker (2008) who state "awareness (through attention) is necessary for noticing which in turn is essential for learning" (p. 248). To sum up, it can be interpreted that the computer is able to provide the visual stimulus to trigger attention and awareness leading to noticing.

5.3.2 Noticing

Noticing is taken to be the impetus that set the motion for language development (Schmitt, 1990; Gass & Mackey, 2006; Gass & Selinker, 2008). According to Schmidt (1995) the Noticing hypothesis asserts that what learners notice in input is what becomes intake for learning. On top of that, the noticing can be deliberate or unintentional but if it is noticed and it becomes intake then noticing is a necessary condition for learning. Gass and Mackey (2006) emphasize the role of noticing in their model in SLA.

They describe the model on interaction and learning with a particular focus on noticing. In their model, during interaction in a conversation, learners receive feedback on their production, which draws attention to the linguistic problems which had arisen

during the interactions. This would subsequently lead the learner to notice gaps between their own production and the target language.

Indeed, the computer screen is the focal point where visual interaction takes place between the user and computer. By default, the screen becomes the first point of contact between the user and computer. Therefore, a kind of "noticing" is already in place well before the actual interactions with the computer begin. The interactions with the computer software or applications inherently facilitate noticing. In addition, the visual nature of many CALL applications tends to promote noticing (Youngs, Ducate & Arnold, 2011).

This study which is designed in a CALL context, noticing had also occurred in the following manner. Specifically, the learners have interacted with the gloss and had obtained modified input which can be regarded as positive feedback in the form of meanings of the unknown words in the form of glosses. This drew the learners' attention to the meanings as given by the gloss against their own initial perceived meanings.

This then allowed them to notice the gap and enhance their vocabulary knowledge of the unknown words. Therefore, it appeared that the enhanced input of the target words had managed to trigger noticing among the students. Before the notion of noticing is proclaimed as a positive attribute to learning, it has to be dealt with carefully. Schmidt (1990) cautions that noticing itself does not result in acquisition, but it is an important starting point for language learning. Schmidt also claims that there are certain factors which can influence noticing in the input. Among them are instruction, frequency, perceptual salience, instructional strategies, learner's processing ability, skill level, students' readiness to notice items and task demands. How these factors are factored into the study is explained next.

As for frequency of encounter, the students only encountered the target words once in the text. However, repeated encounters were possible when and if the students clicked on the words to access the glosses. Also, it has to be pointed out that the students would have also encountered the words in the perceived vocabulary knowledge pre-test as well as the receptive and productive tests in the study, thus increasing the encounter with the target words. Judging from the number of clicks it was apparent that the frequency of meeting the words did not offer much in terms of learning to certain groups of students. For instance, the mid proficiency students who had the most number of clicks lost out the most in vocabulary knowledge compared to other proficiency level students.

In this research, the highlighted words were intended to invoke perceptual saliency and the gloss when activated by learners' computer clicks should provide the opportunity for learning by providing help for vocabulary meaning. In this matter, the saliency of the gloss had probably attracted the students to click on the target words, triggering the initial step in the sequence of language learning in terms of readying the students to notice items. It is clear that the glossed targeted words provided visual saliency perhaps leading to perceptual saliency for the students.

The point on instructional strategy was not incorporated in the study as the current study examines vocabulary development from the learners' interactions with the glosses. As a result no instructional pointers were given for the learners to focus on the targeted words for learning. On the other hand, the learners were told to read the text for comprehension. As for task demands, the task in this study is reading for comprehension that would require students to know the meanings of the words so as to increase their overall comprehension of the text. No other tasks were set in the study.

This lack of tasks could have attributed to the loss of vocabulary knowledge amongst the students in this research. More tasks in the research would have allowed for more processing within the learners which may lead to more reinforced learning.

Finally, skill level and processing ability refer to a learner's ability to attend to both form and meaning in L2 processing. It can be discussed here that from the data obtained it showed that the high proficiency learners could have been better input processors where they were able to change the input into intake more efficiently than the mid and low proficiency levels learners. What can be speculated here is that high proficiency students tend to have better skills and processing abilities. This led them to gain more in the interactions with the glosses.

The question that is unanswered is how to enhance the learners' skills and processing abilities to take advantage of computer-aided glosses or any other aid in language learning? The answer may lie in instructional strategies employed in the classroom which may then be supplemented by glosses or other learning aids. This highlights the case for interaction, be it in the form of learner-computer or classroom interactions. Given the prominence of interaction, this discussion focuses on it in the subsequent section.

5.3.3 Interaction

Interaction is an important factor in language learning. It is clear that second languages are learned through interaction (Long, 1996; Pica, 1994). In the Interaction Hypothesis, Long (1983, 1996) claims that modified interaction can bring about better comprehension and acquisition through the interactional modifications that occur between interlocutors. However according to Warschauer (1999) interaction in itself is not adequate for teenage and adult learners to learn a language.

He argues that finding the ideal balance between form and meaning is a challenge for language teachers. One possible way is to maintain overall communication and meaning, but at the same attend to form. He further emphasises that this is difficult to achieve in conversations as interactions in such instances are quick-paced and it is unlikely for the interlocutors to halt their interactions to attend to form. This astute observation puts the computer learning environment in focus as CALL is able to offer more control to the users or learners, in the sense that the learners can pace and make choices of the computer applications which are relevant and beneficial to them.

It has been acknowledged that within the Input-Interaction framework for SLA in a face-to-face context, interaction can bring about change to learners' interlanguage (Gass, 1997). The interactions in that model in the oral context involve elements such as comprehension checks, clarifications, recasts, and repetitions between the interlocutors. Clearly, the interactions in an oral context are more varied and targeted to solve specific language related deficiencies between the interlocutors. In addition, the crucial factor of feedback comes into play in such interactions, where interactions take place and are even modified further to improve the communication process, hence providing more opportunities for improved comprehension and learning.

This present study offers evidence that within the Input-Interaction framework in a computer environment, vocabulary knowledge development can occur among learners when they obtain modified input. Strictly speaking in this study, the interaction that was provided was one that Chapelle (2003) had specified as learner-computer interaction and not interactionally modified interactions. This is where learner-computer interaction can be useful where the learner has the opportunity and perhaps more time to notice form in the input compared to a face-to-face interactional situation.

The study revealed that these interactions with the modified input in the form of text-only glosses can promote perceived vocabulary expansion as well as productive and receptive types of knowledge in the short-term. Although, the interaction of the learner and the glosses may seem restricted as compared to the interactions that can occur in a face-to-face conversation where more forms of interactions can take place, learning is still affordable even with a confined set of interactional factors (word/sentence; L1/L2 glosses).

Granted the interactional factors are narrow but only by limiting the factors that the researcher was able to clearly identify which factors were more beneficial than the others and how they interplayed to aid vocabulary expansion amongst the learners with different proficiency levels.

Learner-computer interaction may also offer fewer distractions to the learner as compared to a face-to-face context as the learner is focussed on the computer and in a "ready" mode for noticing and interactional possibilities or what Gass' (1997) terms as interactional mode. Therefore, it would appear that a learner-computer interface lends itself as a conducive conduit for learning. The nature of the computer too may prime the students for learning with the students giving their attention to the computer screen where it has been previously discussed that attention can promote noticing and facilitate learning. Gass (1997) deliberates that what is necessary for interaction is that the input has to be made salient. What is crucial is that the input has to be made available for attentional resources to be focused on form or meaning. She continues, "when learners are in an active interactional mode, they can focus on what is necessary for them – their own attention can drive the interaction" (p. 129). It can be seen from the interactions with the glosses in this study that the learners were already in the interactional mode. Interactional mode here may refer to the computer with its screen and mouse interface inherently readies the students to begin clicking and interacting with the computer.

This is evidenced in this study by the clicking done by the learners. Therefore, other elements such as proficiency level or the need to consult the gloss may further drive or thwart the interactions. This was probably seen in the different ways the learners in different proficiency levels interacted with the glosses and benefitted from the interactions differently.

In essence, it means that if there are glosses available for learners which have been made salient, there is a tendency for the learners to access them, creating what Gass said as interactional mode. What increases or decreases the use of gloss would be other variables such as the proficiency level of the learner and their "need" for the glosses.

5.3.4 Input-Interaction model: SLA and CALL learning contexts

In face-to-face second language situations when learners interacted with more competent speakers, the learners were able to notice the gaps or the deficiencies in their language. Their attention was thus directed to these deficiencies which would lead them to notice and take adequate steps to counter these gaps; among them could be requesting for clarification or rephrasing from the more eloquent interlocutor so that language learning can occur. Thus, the initial step in learning began with the learners noticing the gaps in their language.

Within the computer environment, however, the computer acted as the more able interlocutor and one way learners can proceed to obtain feedback was to click on glossed words as in the case of this study. In the computer context, making language forms salient was technically easy and effective.

One of the ways of markedness or making the items salient in a computer text was by highlighting the item differently from the rest of the text, for example, by using a

different font or colour from the rest of the text. Hence, in the learner-computer interface, noticing was externally executed when a text was visually enhanced in the above mentioned manner, thus activating the Noticing Hypothesis by the act of clicking indicating that there was awareness, attention and noticing on the part of the learner.

The use of the gloss in this study was triggered by clicking on the unknown or unfamiliar word which can be interpreted to mean that the learner had noticed a gap in his knowledge and attended to the linguistic form by clicking on the unknown word. Chapelle (2003) acknowledges this as "intra-person" interaction. This may mean that the learner was "attending" to the word and its embedded meaning as provided by the gloss. Another dimension to this was that learners may click on words not so much as a sign for comprehension breakdown or other linguistic deficiencies but could be a sign of curiosity, or was drawn to the technology attached to it, for example, a video clip or an auditory media rather than for its usefulness. What this means was there can be other factors which were drawing the learners' attention and subsequent noticing to the words.

In SLA, learners interacted with their interlocutors to repair a communication breakdown so that the communication process between them was carried on. It was in these interactions, through a myriad process of negotiation of meaning that may facilitate language learning. In a CALL environment such interactions can also be afforded by the computer, either it being the medium of the interaction with other learners or the computer itself providing the platform for the interactions.

The focal point in this study, however, was to examine way how interaction with the computer can lead to vocabulary knowledge development by obtaining modified input in the form of textual glosses. It was the interaction with the glosses which embodies noticing and obtaining modified input through interaction which may lead to vocabulary knowledge increase.

This is supported by Chapelle (2005) who writes that "learners make essential form-meaning connections in learner-computer interaction for instance during reading if the learner stops the input to request for help in the form of vocabulary aid, rephrasing or text transcripts" (p. 55).

In this study, the learners interacted with the glosses to get the meanings of the words which were unfamiliar to them. Interaction was therefore seen as the source of learning or in this case the platform to develop learners' vocabulary knowledge. This is in line to Ellis (2012) who elaborates that evidence for learning does not come from the interaction itself, but independently from the tests or tasks that precede and follow the interaction, reflecting the fact that "interaction is seen as the source, not context, of learning" (p. 240). The vocabulary tests in the study provided evidence of vocabulary expansion or otherwise.

The interactions between the learners and the gloss can be described as a "highly structured form of negotiation of meaning as learners control the interaction in its various modified forms" (Chapelle, 2007, p. 103). In this sense, when there was a breakdown in reading comprehension, learners were able to request for help from the glosses. This request for help is seen as a type of negotiation of meaning that occurs in the interaction of the learners in conversation-like situations. Smith (2004) states that the negotiation that triggers interactional adjustments which facilitates acquisition as it connects input, internal learner mechanisms and finally producing output. This piece of evidence supports that input modifications and interaction are important for second language acquisition. Thus, the results of this study have shown that interaction with the textual glosses can aid language acquisition, in this case is vocabulary development.

In a SLA context, output serves as the evidence of learning. Output is the language which is produced by the learners either in speech or writing. In this research output from the learners was their performance in the vocabulary knowledge tests.

Pointedly, the manifestation of the Comprehensible Output Hypothesis was in the form of perceived, receptive and production tests to measure learners' vocabulary knowledge after interacting with the glosses. The assumption was the vocabulary knowledge test scores reflected that the vocabulary knowledge of the learners.

This would mean that the learners' vocabulary knowledge had been through the intake and integration stages of the Input-Interaction model corresponding to the immediate and delayed vocabulary tests scores. The vocabulary tests were a direct measure of the vocabulary knowledge. What was learnt from this kind of testing that it was "untimely" in the process of learning to measure output. What was needed was more "pushed output" type of tasks to be in place before this kind of direct testing of vocabulary knowledge can be administered. Also, with such tasks, the students would have more opportunities to notice their errors or to be given feedback on their output. Besides, more tasks or probably vocabulary practice should be in place before direct testing is conducted. Meaningful vocabulary tasks too would be a more reliable measure of vocabulary knowledge, instead of vocabulary tests as used in this research.

To make clearer the interface between input and interaction in the two environments: face-to-face and computer, Table 5.1 maps out the input-interaction factors in a face-to-face or conversational framework to a CALL context as found in this study. The differences are segmented into input, interactional features, obtaining modified input and the nature of interactions.

Table 5.1: The mapping of input-interaction factors from an oral to a computer context as in the study

Face-to-face context	Computer
Input is modified by simplification of vocabulary, syntax or discourse	Online input is modified by simplifying online text by taking into account length, sentence structure, and grammar. Target vocabulary is modified by providing meaning at word and sentence levels.
Interactional modifications comprise wide negotiation features such as clarification, comprehension checks, repetition etc.	Interactional modifications are restrictive to access to gloss conditions
Obtaining modified input from the negotiation process	Obtaining modified input from the modified glosses
Interaction is spontaneous and mostly unplanned.	Interaction is measured and planned.

5.4 Implications to research

The research undertaken had initially piloted the gloss use by presenting the whole range of glosses to the learners. What is meant is here that the whole range of glosses were made available to the learners, this is, WBM, SBM, WEN, SEN. The learners then chose the type of gloss which they preferred to interact with. The results from the pilot test showed that no clear pattern emerged from this kind of gloss use. As a result of this pilot testing, the final experiment was carried out by randomising the learners into specific gloss conditions: WBM, SBM, WEN, and SEN. By doing so, the trade off was the choice of the learners in choosing the type of gloss that they preferred is lost. This would have been valuable data as it would have tapped into the preference of the students with the type and language of glosses. On the positive, however, what was documented was the efficacy of the type and language of gloss was obtained.

In doing this, there were some patterns which have emerged from the interactions with the glosses.

In this study, the 13 words which had been glossed in the text had a few ramifications on the study. Firstly, more words in the text would have probably revealed more definite patterns of clicking and interactions with the glosses. Secondly, it had also affected the vocabulary tests. Although care was taken in the study on the time interval of the immediate test and the subsequent delayed test, more words in tests would have ensured a higher reliability as the test effects would not have been apparent. However, given the length of the text, it was not possible to have more than 13 words glossed in the text.

The design of the research where only the perceived vocabulary test was given before the experiment should also be re-examined. The vocabulary receptive and productive tests were given only as post-tests (immediate and delayed). This was done because of the low number of words in the research hence, there was concern of test effects which would affect the results of the receptive and production tests if they were also given prior to the experiment. Furthermore, this design was in line with an early interaction study done by Ellis and He (1999) where they too had administered only post-tests to measure interactional effects. Nonetheless, in this study if both the tests were administered as pre-tests, perhaps the vocabulary development of the students could have been traced more meaningfully, provided the test effects of the tests could be controlled. In total, studies which investigated interactions with words to enhance vocabulary knowledge or learning have to be careful of the number of words used. Indeed, in this study, the small number of words and the placement of the productive and receptive vocabulary tests had affected the research design and to a certain extent the outcomes of the research.

Another implication on research is that there should be other measures that investigated learner-computer interactions from all possible angles. Firstly, an attitude questionnaire could have provided the learners' impression of the interactions which they had taken part in. Secondly, think-aloud protocols (TAPs) could also be utilized the show the pathways of the learners' thinking processes while they interacted with the computer. All these can complement the process data obtained from tracking devices in the computer.

The results also showed that learners gained more productive vocabulary knowledge from the interactions as compared to receptive type of knowledge. This can be attributed to the sentence type of glosses which provided more context to the learners, hence they were able to make gains. The other possibility was that the test to measure productive vocabulary knowledge may have been too simple for the learners. On the other hand, less gain in receptive vocabulary knowledge was recorded and this could be attributed to the multiple choice format of the test measuring this knowledge. The test could have been difficult for the students. On the other hand, the productive vocabulary knowledge test which required students fill in the gaps with the aid of initial letters of the target word could have been too easy or they could have guessed the answer because of the given initial letter. Therefore, a more reliable way of measuring both productive and receptive vocabulary knowledge in research is needed. It could be in the form of multiple tasks where aspects of both types of knowledge can be measured more meaningfully and accurately (Please see section 5.6.2 for a further discussion on receptive and productive vocabulary knowledge).

5.5 Implications to pedagogy

The discussion on implications to pedagogy is divided into two. Firstly, the overall implications of using glosses as a computer aid are put forth. Secondly, a more precise discussion on the implications of gloss use to vocabulary learning including types of vocabulary knowledge is given.

Textual glossing in CALL as in this study may enhance vocabulary learning in the short-term. It would seem that learning vocabulary through computer-textual glosses can be an interim solution to the teaching and learning of vocabulary, in particular low frequency words. Glossing of unfamiliar or difficult words in CALL can act as an autonomous vocabulary episode for the students to be later complemented with direct teaching. It can complement direct teaching as direct vocabulary teaching is time consuming and as Parry (1993) notes that "it is simply not economic to spend precious minutes on items whose chances of reoccurrence may be low" (p. 2).

From the research, albeit cautiously, the researcher sums up that interactions with unfamiliar words which were glossed in word and sentence level in English or learners' L1 may benefit the learners' vocabulary knowledge in the short term. For retention of knowledge however, more robust and "pushed output" tasks have to be designed to sustain knowledge from the initial gain made by textual-only interactions. The tests measuring the productive and receptive knowledge in this study did not allow the learners sufficient opportunities for them to produce output. Therefore for more long-term gains by teaching, more tasks are designed to create more instances for noticing, interaction and processing which would enable favourable long-term gains in vocabulary knowledge.

The results of this research also pointed out that students in different proficiency levels gained differently from the gloss use. Learners with different proficiency levels will have a variance in their gains of vocabulary knowledge. High proficiency level students may be able to learn vocabulary more independently with minimal direct teaching intervention compared to low and medium proficiency level students. Besides that, teachers should be aware that probably low and mid proficiency level students need more sentence and L1 type of glosses, while higher proficiency level students may gain from sentence and L2 glosses.

It also appeared that mid proficiency learners were the ones who although showed a higher amount of interaction (from the number of clicks) made losses which were higher compared to the low and high proficiency levels. This was attributed to their "click happy" or "surfer" behaviour leading to shallow processing of the vocabulary items. The reasons for such superficial processing certainly need research which taps into the processing facility of the learners with different proficiency levels. Hence, it is important for language teachers to be aware of this when designing gloss for students in their classes to maximise the benefits of gloss use.

Picking up this point of look-up behaviour, the researcher suggests that when dealing with CALL, teachers should also guide students on their look-up behaviour. For instance, teachers should encourage students to develop more "searcher" kind of look-up behaviour instead of "click-happy" or "surfer" kind of look-up. It would be useful for teachers to inculcate a more "searcher" type of look-up behaviour when dealing with computer aids such as glosses. The study suggests that there may be an inherent pattern linking these two variables of look-up behaviour and proficiency.

From the perspective of testing, teachers should also bear in mind that the type of vocabulary tests will affect the type of vocabulary knowledge tested. For instance, the results from this study indicated that students did better in gap-filling, sentence level format which tested productive vocabulary knowledge than multiple-choice type of test for receptive knowledge of the target words.

If computer textual-glossing in an online reading context is used to teach vocabulary, it is best for teachers to be explicit in their instructions whether the lesson is for reading comprehension or vocabulary learning. This would in a way direct the students for the learning activity ahead and their attention can be channelled appropriately – either reading for meaning or form. The next section of the discussion deals with the use of glosses on vocabulary learning.

5.5.1 Vocabulary learning

For language learning to occur, there has to be conditions to be met by the language learner. From the literature on learning vocabulary, Schmitt (2002) writes that the conditions for vocabulary learning is firstly, input should contain a small number of unknown words, secondly, there has to be a large quantity of input. Thirdly and more relevant to this thesis is his affirmation that:

> learning will be increased if there is more deliberate attention to the unknown vocabulary through the occurrence of the same vocabulary and through consciousness-raising of unknown words as they occur through glossing, dictionary use and highlighting in the text. (p. 40).

In relation to this study, the first condition was met as only 13 words were used as target glossed words in the text. The second condition was not met as research constraints such as time and methodological issues made it difficult for a large quantity of input to be made available to the students. The third condition was of course met as the words in the text were glossed and highlighted in the text. This meant that this study had set a favourable condition for vocabulary development.

The input which had been enhanced opens up another issue of whether the learning that followed the interaction was incidental or intentional. Before this topic is discussed further, it is important to define the terms incidental and intentional learning. In order to define incidental learning, it is best contrasted to intentional learning. In incidental learning, learners are required to perform some information processing task without being told beforehand that they will be tested later in their recall of that information. In direct contrast, in intentional learning conditions, learners are told in advance that their recall would be tested after the completion of the task.

In her discussion on incidental learning, Gass (1999) describes it as "a by-product of other cognitive exercises involving comprehension" (p. 319). In the same paper, she points out that the notion of attention is also available in incidental learning. She argues that incidental learning can occur in both situations, that is, with and without attention. A completely incidentally situation exists when no exposure has been given to learners on the target items to be learnt. In other instances, there is some kind of exposure, although the learning is not intended.

In the context of this study, the latter condition of incidental learning applies as there was exposure of the vocabulary items made to the learners in the online text and glosses. The words were made salient by highlighting them in a different colour from the rest of the text. However, the learners were not instructed on learning vocabulary through the highlighted words and gloss.

What is meant here is that at the outset of the experiment, the students were told to read the online text for comprehension. Although a reading comprehension task is included, the task is to deflect the students from perceiving that the study is on an inquiry into gloss use and vocabulary development. Moreover, they were told that unfamiliar words in the text were highlighted and there was a glossary attached to the words which can be viewed by clicking on the target words. In sum, they were not told that the whole exercise was on vocabulary learning. The focus is on comprehension of the text and not on intentional vocabulary learning. Although this study has contrasted the differences between incidental and intentional learning, the stance taken in this research is that what is more important is the quality of processing than the type of learning as claimed by Gass and Mackey (2012).

From the literature of this study, it is seen that vocabulary learning can be mapped onto two particular approaches. One is Nation's (1990) vocabulary learning theory which involves the elements of Noticing, Retrieval and Generative. On the other hand is Laufer and Hultsjin's (2001) Involvement Load theory which essentially posits that the more the learner is involved with the vocabulary, there is more processing of the vocabulary items which in turn may lead to vocabulary learning.

Related to the use of glosses as in this study, both the approaches are applicable for vocabulary learning. For instance, in keeping to Nation's approach, the first element is met when students notices the gap in their linguistic knowledge, which prompted them to retrieve the meaning. In this study, it would be the clicking and the interaction with the glosses. Finally, in the generative stage, the students generated the vocabulary items in the vocabulary tests. In the study, the vocabulary tests were in the form that matched the generating strategies in Nation's vocabulary learning theory, however, the tests were more placed for production instead of generating strategies.

As for the Involvement Load theory, the three major components of the theory were also present in this study. Firstly, the students read the text and met the unfamiliar words which triggered a need for them to understand the words. Secondly, there was search in the form of clicking on the target words to access the glosses, and thirdly, the evaluation stage where the students evaluated the appropriacy of words for use in the vocabulary tests.

Both the approaches described above are applicable in this study. However, it appears that Nation's Vocabulary Learning theory is more suited to explain the vocabulary learning in this study with the elements of Noticing, Retrieval and Generative, in tandem with noticing, interaction and output as manifested in this study. As for Hultsjin's Involvement Load theory it may not be applicable here because the involvement load or processing of the students with the target vocabulary is insufficient for effective learning. It can be assumed that the interactions with the textual gloss and the subsequent vocabulary tests in the study did not provide enough involvement for the students. Furthermore, the study did not attempt to document the level of processing of the students.

5.5.2 Vocabulary knowledge: Perceived, Receptive and Productive

The literature on vocabulary learning shows evidence that receptive knowledge is more readily gained when compared to productive knowledge. It is documented that the ideal vocabulary learning cline should move from a receptive stage to more productive use of vocabulary (Laufer & Paribakht 1998; Laufer & Goldstein, 2004; Ortega, 2009). In simple terms, it means learners can recognise more words than they can actually use. However, this trend of development was not seen in this study.

From the data, it can be seen that students gained more productive vocabulary knowledge compared to receptive knowledge. This mismatch could have resulted from the tests used in the study. What could have transpired resulted from the nature of the tests used in the study. One probable reason was that the productive test with its gap-filling and initial letters format could have given more readily clues of the right answers to the students. The next reason was that the sentence-level format may have provided the necessary context for the students to arrive at the right answer, thus they were able to score higher in the productive test. It could also be reasoned out that the lenient scoring guide could also be the reason the students did better in the productive test. The scoring guide allowed for marks to be awarded even though the inflections were not correct. In contrast, the multiple choice format of the test measuring receptive vocabulary knowledge without any context may have been difficult for the students. They were unable to provide the correct meaning of the target words. Hence the disparity in performance of the students in these two tests. The conclusion that can be made is that the design and format of the tests could play a significant role in the outcome of the tests.

There is also a need to discuss the nature of the perceived vocabulary knowledge of the students to actual receptive and productive vocabulary knowledge. The findings revealed that the students had perceived they knew the meanings of the targets words when it fact, they did not. This was clear when the vocabulary test scores of the perceived test was compared to the actual performance in the receptive and productive vocabulary knowledge tests. The students in all proficiency levels obtained high scores for the immediate perceived vocabulary knowledge test which was administered after the gloss use. What this suggests is that the students' perception of their own knowledge was higher than what they possess.

In terms of the gloss use, it may have given the learners a false sense of "learning" the target vocabulary which did not translate into actual gains as seen in the receptive and productive vocabulary test scores. Alternatively, the glosses may have provided confidence to the students who then perceived that their knowledge of the target vocabulary items had increased after using the glosses. This is evident in the results of the tests that were conducted immediately after the gloss use. This can be useful as a classroom strategy to build on this confidence by providing direct instruction or meaningful tasks to further develop more concrete vocabulary knowledge instead of perceived vocabulary knowledge.

5.6 Implications to CALL

From the perspective of designing reading materials for readers, the findings of this study may suggest that for difficult or unfamiliar words, they can be glossed in the students' L1 or L2, at either word or sentence level as this may allow the readers to gain knowledge of the words immediately to improve their vocabulary knowledge.

It has been indicated that CALL is able to provide the necessary element of saliency which can act as the impetus to language development. More importantly, this research has demonstrated that more direct and straight forward type of glosses can help students. The nature of such interactions is probably sufficient if the aim of the webpage designer is for the reader to develop short-term vocabulary knowledge gains. This translates into simpler and more time-saving efforts in creating computer glosses.

Another implication to CALL is to look at the other end of the input-output spectrum, in other words the output. The study had looked at input from the computer in the form of textual glosses and interaction of the students and the gloss. Output, on the other hand, was measured by the vocabulary tests.

From the perspective of CALL, it is possible too for the computer to play a role in the output in the sense that the computer is able to provide opportunities for students in producing output. In these instances, there would be more room to enable the students to notice their linguistic gaps or errors and make more suitable and correct adjustments to their output to benefit their learning.

5.7 Limitations and recommendations for future research

Among the limitations to the study is of course, the study was not able to investigate the internal working of the learners' minds. The only observable element was the outward clicking of the students which could have reflected awareness, attention and noticing. Even within clicking, this study was unable to distinguish between clicking and interaction. Perhaps, in future studies tracking time taken for a click and interaction is defined which may then act as a delineator between the two notions. However, before that can occur, the issue of how much time constitutes a click or interaction has to be explored further.

It is suggested that future research on learner-computer interactions in the context of vocabulary learning has to weigh in the number of words and text length in order to obtain more distinct pattern in interactional patterns. This present study as well as previous studies such as Yoshii (2006) and Laufer & Hill (2000) which have used short texts and a small number of words and have shown promising significant findings on interactions. Nevertheless, the findings could be more meaningful in terms of the generalizibility of the findings on the interactions for vocabulary learning. What this means is that probably a small number of words can be used for investigating vocabulary learning. However, if interactions are the focus of the study, then a larger number of words is required.

Most research in CALL had utilized non-authentic texts. It is suggested that future research in glossing should use authentic texts which are readily available on the Internet. The use of authentic text in research is reinforced by Plass and Jones (2005) who state:

> existing research has primarily been conducted with materials solely designed for language teaching, leaving open the question as to whether the findings obtained with these materials transfer to materials that were not specifically designed for learners of the language, such as target culture Web sites and videos that provide authentic language experiences to students. (p. 479).

Another plus point for the use of an authentic material for research is that authentic material will provide the students with materials beyond the learners' linguistic level, hence challenging students to read what would be happening around the learners in real-life situations (Gettys, Imhof & Kautz, 2001). For future studies, it is also recommended that perhaps more than one text is used as this could also heighten the knowledge of different types of texts (input) that affect interaction and language learning. It is also recommended that more words are glossed which can probably detect a clearer pattern of gloss use.

One other aspect that future research can look into is the types of user interface in learner-computer interactions. Thus far, clicking is the common interactive method of interface with the computer. Other methods of interactions such as flipping or zooming may display different types of patterns of interaction and outcomes for learning (Oh, Robinson & Lee, 2013).

Lastly, the study had only utilized students from one higher learning institution, comprising only bumiputera students. As a result, the findings of this study cannot be generalised to a larger population.

Furthermore, the use of *Bahasa Melayu* as one of the languages in the glosses restricted its usage to students who know the language, thus limiting its generalizibility.

5.8 Conclusion

This final section begins by revisiting the main findings from the study to derive the conclusions of the research. It also highlights what can be learnt from this research to the theoretical model of SLA to a CALL context, in particular the Input-Interaction model which was put forth in Chapter 1.

The conclusion that can be made thus far is that the clicking behaviour of the students with the glosses did not differ significantly among the students of different proficiency levels.

As for specific interactions, textual type of glosses, word and sentence, in L1 and L2 can help learners expand their vocabulary knowledge. This means that word or sentence glosses in either L1 or L2 were able to help the students develop their vocabulary knowledge in the short-term. In total, direct, simple and straightforward textual glosses as utilised in this study have their value in vocabulary development t with some caveats. It can be attested that textual glosses, word and sentence type in BM and English can develop students' vocabulary knowledge. However, the interactions with the glosses were unable to make the learners retain some types of vocabulary knowledge for students in certain proficiency levels as in the low and mid proficiency levels.

Still on the issue of the type of glosses, sentence type of glosses benefitted the learners in all proficiency levels except for the initial perceived vocabulary knowledge and receptive vocabulary knowledge. The benefit of the sentence type of glosses can be attributed to more context which was provided in the glosses.

In support of this were previous studies (Hulstijn, 1993; Knight, 1994; Chun & Plass, 1996; Grace, 1998; Lomicka, 1998; Nagata, 1999; Laufer & Hill, 2000; and Yoshii, 2006) which have shown the value of context in sentence type of glosses.

In contrast, it was seen that word type of glosses had helped the learners gain receptive type of knowledge. What this may mean is that both types of glosses can generally be used for vocabulary learning, however, care and thought should be taken when designing such glosses in relation to the type of vocabulary knowledge that is being taught. Therefore what can be learnt from this study is that sentence type of glosses may aid productive vocabulary knowledge, while word glosses help receptive vocabulary knowledge.

For patterns of interactions in relation to proficiency level, it is evident that low and mid proficiency levels have similar patterns. Closer observation of the data of low and mid proficiency levels showed that the interactions with sentence and L1 glosses aided them for perceived and productive vocabulary knowledge, while L1 and word type glosses aided them for receptive vocabulary knowledge. On the other hand, high proficiency level learners benefitted from interactions with L2 glosses in all types of vocabulary knowledge. The interaction patterns of the students in different proficiency levels differed together with the benefits it brought to the students' vocabulary knowledge. It can be concluded that there is a need for more varying types of interactions which can help the learners. What can be speculated here is there has to be more varied multiple exposures to the gloss so that vocabulary knowledge can be sustained. The reason behind this is perhaps with more varied kind of glosses, there would be more processing, leading to knowledge being held in the long-term memory of the learners which can then be integrated into the learners' interlanguage.

As for proficiency level, it is important to keep in mind that learner proficiency is factor to be seriously considered in CALL interactions and vocabulary development. The study had in a limited way revealed that high proficiency level students gained the most in the use of glosses for vocabulary development, followed by mid and low proficiency level students.

From the point of theory, it can be agreed that the Interaction Approach is an approach which is important, relevant and effective for SLA, not only in face-to-face kind of interactions but also in a CALL environment. This study has reinforced that learner-computer interactions in the form of glosses is beneficial for language learning. It was also demonstrated that the students' engagement with the textual-only computer glosses in L1 and L2 can help in developing the students' vocabulary knowledge. Further, the Input-Interaction model in a conversational setting can be extended to a CALL environment with some constraints. The different hypotheses that interplay in the model had some ramifications in the CALL setting.

Firstly, it appeared that in CALL, the Noticing Hypothesis was easily stimulated by making the glossed words salient in a different colour from the text which meant that noticing was induced on the part of the learners externally. The students who interacted with the glosses were aware of the glossed words by the text enhancement. This meant that noticing was not internal but externally induced.

Noticing externally may encourage interactions with the gloss but it may not be as effective as internal noticing. It is also necessary to note that the tests in the study had probably provided some type of initial exposure leading to noticing of the target words. Nonetheless, that kind of noticing was probably not strong enough to trigger awareness and processing of the deficient vocabulary knowledge. Although this type of noticing together with the externally-induced noticing by the glosses of the target words, it is clear that such noticing and the exposure to the glosses did not provide

sufficient processing for more sustained language development to occur. On a cautionary note, although the benefits of noticing to learning seem attractive, it is also important to realise that the notions of internally and externally-induced noticing are difficult to distinguish. This brings up the point that any outcome from this notion of noticing has to be dealt with care.

Secondly is the point which is related to the Frequency Hypothesis in terms of the exposure to the glosses. In the Frequency Hypothesis (Ellis, 2002) it is stated that learning can be increased if there are more frequent exposures to the items to be learnt. On the other hand, infrequent items can also draw attention to the learner as seen in studies by Chun & Plass (1996) and Laufer & Hill (2000). Therefore, it is difficult to pin down the number of clicks as in the study to the Frequency Hypothesis. There are other factors to consider such as Noticing, interaction and deep processing in tandem with the Frequency Hypothesis. Although the glosses offered repeated exposures to the target words, these kinds of repeated measures were not necessarily helpful. The students did not seem to interact with the glosses repeatedly in a manner which can be helpful for them. More frequent interactions with the glosses could have provided students with the needed engagement with the glosses for more long-term learning.

Thirdly, as for the notion of interaction as espoused in the Interaction Hypothesis, it is clear from the study that interactions with the textual glosses in both L1 and L2 can benefit the development of vocabulary knowledge Nonetheless, other variables such as learner proficiency, processing ability and the readiness to learn are in place. The variables of processing ability and the readiness to learn are individual characteristics and are obviously not uniform amongst learners. These characteristics are not tangible and cannot be taught to the learners. What is suggested is that the learners' awareness of these characteristics can be heightened by language teachers.

It can be concluded that textual glosses In L1 and L2 can help in vocabulary learning. However, in maintaining vocabulary knowledge these learner variables have to be factored into the use of glosses as a whole.

Next, from the point of the Comprehensible Output Hypothesis, output in this investigation was tested through the vocabulary tests. It is learnt that the interactions with the glosses alone were inadequate for more sustained vocabulary learning. Additional tasks that provide more exposure and practice of the vocabulary items to the learners should be in place before testing output. This meant that the timing of such tests as administered in the study was premature. Meaningful vocabulary tasks should be placed after the interactions which could lead to more vigorous and prolonged vocabulary development. This knowledge could then be measured by direct vocabulary tests if required.

The closing remark on the theoretical model is that Gass' (1997) Input-Interaction model can be extended to a CALL environment with restraint. It can be utilized for language learning, in particular, computer-learner textual gloss interactions are a viable platform for short-term language learning. However, the interactions that the computer provides may not be as wide-ranging as in a face-to-face context. The textual glosses, although meaningful, may have provided only superficial interactions which did not lead to deeper processing which is needed for long-term language learning. This could also mean that the textual glosses in the study did not reveal the expected benefits of glosses to the students. In other words, the interactions need to be more varied and this where computer technology has to be expanded so that it can mimic real-life exchanges to help in more prolonged language learning. In short, learner-computer interactions have to mirror interactions in the oral context to benefit the learners. Notwithstanding, the early statements about the overuse of technology made in the Introduction Chapter have to be borne in mind.

Furthermore, the use of technology has to be in line with SLA theories as its use is meaningless, if it is not aligned to theory which can explain language learning.

Going back to one of the assumptions of the study, where it was stated that internal working of the minds of the language learner is important for learning to occur. Although a CALL environment can provide useful learning settings, finally, it is the innate mechanisms that will impact how the learner utilises the provisions for language learning be in a computer environment. This calls for CALL to have more innovative ways in technology that can supplement, complement and heighten the workings of the learners' minds. This forefronts the human element in language learning bringing up the premise that although the CALL environment can bring about language development, the human fundamentals in language learning that is, the learners' minds and interactions are indispensable to successful language learning.

To end this thesis, it is envisaged that how these human elements interact, engage and integrate with technology is the way forward for language learning. It is hoped the input provided by thesis which has indicated that word and sentence glosses in L1 and L2 can help students develop their vocabulary knowledge may yield useful contributions to this already intense area of computer gloss use. Collectively, it will provide the area with breadth and depth where findings from such studies can benefit the major stakeholders in SLA; ultimately the most gains should be reaped by the language learners themselves.

References

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

Al-Seghayer, K. (2001). The effect of multimedia annotation modes on L2 vocabulary acquisition: A comparative study. *Language Learning & Technology*, *5*(1). Retrieved from http://llt.msu.edu/vol5num1/alseghayer/default/html.

Allum, P. (2004). Evaluation of CALL: Initial vocabulary learning. *ReCALL*, *16*(2), 488-501.

Anderson, N.J. (2003). Scrolling, clicking, and reading English: Online reading strategies in a second/foreign language. *The Reading Matrix, 3*(3). Retrieved from citeseerx-ist.psu.edu.

Aweiss, S. (1994). Situated learning in technology: The case of computer-mediated reading supports. *Journal of Educational Technology Systems*, *23*(1), 63-74.

Baddeley, (1986). *Working Memory*. Oxford: Oxford University Press.

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

Bland, S.K., Noblitt, J.S., Armington, S., & Gay, G. (1990). The Naive Lexical Hypothesis: Evidence from Computer-Assisted Language Learning. *The Modern Language Journal*, *74*(4), 440-450.

Chall, J.S. (1983). *Stages* of *reading development*. New York: Harcourt Brace.

Chapelle, C. A. (2001). *Computer applications in second language acquisition.* Cambridge: Cambridge University Press.

Chapelle, C. A. (2003). *English language learning & technology*. Amsterdam: John Benjamin's Publishing Co.

Chapelle, C. A. (2004). Technology and second language learning: expanding methods and agendas. *System*, *32*(4), 593-601.

Chapelle, C.A. (2005). Interactionist SLA theory in CALL research. In J. Egbert & G.M. Petrie (Eds.), *CALL research perspectives* (pp. 53-64). Mahwah, NJ: Lawrence Erlbaum.

Chapelle, C. A. (2007a). CALL in the year 2000: Still in search of research agenda? *Language Learning & Technology, 1*(1), 19-43. Retrieved from http://llt.msu.edu/vol1num1/chapelle/html.

Chapelle, C. A. (2007b). Technology and second language acquisition. *Annual Review of Applied Linguistics*, *27*(1), 98-114.

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

Chun, D. M., & Plass, J. L. (1996). Effects of multimedia annotations on vocabulary acquisition. *The Modern Language Journal*, *80*(2), 183-198.

Chun, D. M., Mayer, R. E., Leutner, D & Plass, J.L. (1998). Supporting visual & verbal learning preferences in a second-language multimedia learning environment. *Journal of Educational Psychology*, *90*(1), 25-36.

Chun, D. M. (2001). L2 reading on the web: Strategies for accessing information in hypermedia. *Computer Assisted Language Learning*, *14*(5), 367-403.

Chun, D. M., & Payne, J. S. (2004). What makes students click: working memory & look-up behaviour. *System*, *32*, 481-503.

Chun, D. M. (2006). CALL technologies for L2 reading. In L. Ducate & N. Arnold (Eds.), *Calling on CALL: From theory & research to new directions in foreign language teaching* San Marcos: CALICO

Chun, D. M. (2011). CALL Technologies for L2 reading post web 2.0. In N. Arnold & L. Ducate (Eds.), *Present and future promises of CALL: From theory and research to new directions in language teaching* (pp. 131-170). Texas: CALICO.

Cohen, J. (1988). *Statistical power analysis for the behavioural sciencies*. Routledge.

Coiro, J. (2003). Exploring literacy on the internet: Reading comprehension on the internet: Expanding our understanding of reading comprehension to encompass new literacies. *The Reading Teacher*, *56*(5), 458-464.

Cook, V. (1999). Going beyond the native speaker in language teaching. *TESL Quarterly, 33*(2), 185-209.

Cook, V. (2001). *Second language teaching and learning*. Arnold, New York.

Cumming, G., Cropp, S., & Sussex, R. (1994). On-line lexical resources for language learner: Assessment of some approaches to word definition. *System, 22*, 369-377.

Daller, H., Milton, J., & Treffers-Daller, J. (2007). *Modelling and assessing vocabulary knowledge*. Cambridge: Cambridge University Press.

Davis, J. & Lyman-Hager, M. (1997). Computer & L2 reading: Student performance, student abilities. *Foreign Language Annals, 30*(1), 58-72.

De la Fuente, M, J. (2003). Negotiation and oral acquisition of L2 vocabulary. *Studies in Second Language Acquisition*, *24*, 81-112.

De Ridder, I. (2002). Visible or invisible links: Does the highlighting of hyperlinks affect incidental vocabulary learning, text comprehension, and the reading process? *Language Learning & Technology*, *6*(1), 123-146. Retrieved from http://llt.msu.edu/vol6num1/DERIDDER/default.html.

Ellis, N. (2002). Frequency effects in language processing: a review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition, 24*, 143-188.

Ellis, R. (1985). *Understanding second language acquisition*. Oxford: Oxford University Press.

Ellis, R. (1997). *Second language acquisition.* Oxford: Oxford University Press.

Ellis, R. (1999). *Learning a second language through interaction*. Amsterdam: John Benjamin's.

Ellis, R., & He, X. (1999). The roles of modified input & output in the incidental acquisition of word meanings. *Studies in Second Language Acquisition, 24*, 223-236.

Ellis, R. (2003). *Task-based language learning and teaching*. Oxford: Oxford University Press.

Ellis, R. (2008). *The Study of second language acquisition*. (2nd Ed.). Oxford: Oxford University Press.

Ellis, R. (2012). *Language teaching research and language pedagogy*. West Sussex: John Wiley & Sons, Inc.

Ercetin, N. (2001). *Second language reading in a hypermedia environment: The role of proficiency, annotation use, text format, & prior knowledge*. (PhD thesis). Retrieved from ProQuest Dissertations & Theses database.

Færch, C., & Haastrup, K. R. Phillipson. (1984). *Learner language & language learning*. Clevedon: Multilingual matters.

Fischer, R. (2007). How do we know what students are actually doing? Monitoring students' behaviour in CALL. *Computer Assisted Language Learning, 20*(5), 409-442.

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

Gass, S.M, & Selinker, L. (1994). *Second language acquisition*: *An introductory course.* New Jersey: Lawrence Erlbaum.

Gass, S. M, & Varonis, E. M. (1994). Input, interaction & second language production. *Studies in Second Language Acquisition, 16,* 283-302.

Gass. S. (1997). *Input, interaction & the second language learner*. New Jersey: Lawrence Erlbaum Associates.

Gass, S. (1999). Incidental vocabulary learning. *Studies in Second Language Acquisition, 21*(2), 319-333.

Gass, S. M., & Mackey, A. (2006). Input, interaction and output: An overview. *AILA review*, *19*(1), 3-17.

Gass, S. M., & Mackey, A. (2007). Input, interaction, and output in second language acquisition. *Theories in second language acquisition*, 175-199.

Gass, S. M., & Selinker, L. (2008). *Second language acquisition. An introductory course*. 3rd Edition. New York: Routledge.

Gass, S. M., & Mackey, A. (2012). *The Routledge handbook of second language acquisition*. Routledge.

Gettys, S., Imhof, L. A. & Kautz, J. O. (2001). Computer-assisted reading: The effect of glossing format on comprehension & vocabulary retention. *Foreign Language Annals, 34*(2), 91-108.

Grabe, W. (1991). "Currents developments in second language reading research". *TESOL Quarterly*, *25*(3), 375-406.

Grabe, W. (2004). Research on teaching reading. *Annual Review of Applied Linguistics*, *24*(1), 44-69.

Grace, C. (1998). Retention of word meanings inferred from context & sentence-level translations: Implications for the design of beginning level CALL software. *The Modern Language Journal*, *82*(4), 533-544.

Groot, P.J.M. (2000). Computer assisted second language vocabulary acquisition. *Language Learning and Technology*, *4*(1), 60-81. Retrieved from http://llt.msu.edu.vol4num1/groot/default.html.

Hatch, E. (1978). Discourse analysis and second language acquisition. In E. Hatch (Ed.), *Second language acquisition: A book of readings* (pp. 401-435). Rowley, MA: Newbury House.

Hatch, E. & Hossein Farhady. (1981). *Research design and statistics for applied linguistics.* Rowley: Newbury House Publishers, Inc.

Hayden, S. (1997). *An investigation into the effect & patterns of usage of a computer-mediated text in reading comprehension in French*. (Unpublished PhD thesis). The Pennsylvania State University.

Haynes, M., & Baker, I. (1993). American and Chinese readers learning from lexical familiarization in English texts. *Second Language Reading and Vocabulary Learning*, 130-152.

Hegelheimer, V. (1998). *Effects of textual glosses & sentence-level audio glosses on on-line reading comprehension & vocabulary recall*. (Unpublished PhD thesis). University of Illinois at Urbana-Champaign.

Hegelheimer, V., & Chapelle, C.A., (2000). Methodological issues in research on learner-computer interactions in CALL. *Language Learning & Technology*, *4*(1), 41-59.

Henriksen, B. (1999). Three dimensions of vocabulary development. *Studies in Second Language Acquisition, 21*, 303-317.

Hornby, A.S. (2000). *Oxford Advanced Learner's English- Malay dictionary.* Shah Alam: Fajar Bakti.

Hornby, A.S. (2005). *Oxford Advanced Learner's Dictionary* (7th ed.). Oxford: Oxford University Press.

Horst, M., & Meara, P. (1999). Test of a model for predicting second language lexical growth through reading. *Canadian Modern Language Review, 56*(2), 308-328.

Hu, M., & Nation, P. (2000). *Unknown vocabulary density and reading comprehension.* Reading in a Foreign Language, *13*(1), 403-430.

Hulstijn, J. H. (1993). When do foreign language readers look up the meaning of unfamiliar words? The influence of task & learner variables. *The Modern Language Journal, 77*(2), 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, & reoccurrence of unknown words. *The Modern Language Journal, 80*(3), 327-339.

Internet word stats. Usage and Population statistics. (n.d.) *Internet Word Stats*. Retrieved from http://www.internetworldstats.com/

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

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

Ko, H. M. (2005). Glosses, comprehension, & strategy use. *Reading in a Foreign Language, 17*(2), 18.

Ko, H. M. (2012). Glossing & second language learning. *TESOL Quarterly, 46* (1), 56-79.

Kost, C.R., Foss, P., & Lenzini, J.J. (1999). Textual & pictorial glosses: Effectiveness of incidental vocabulary growth when reading in a foreign language. *Foreign Language Annals, 32*(1), 89-113.

Krashen, S.D. (1982). *Principles and practices in second language acquisition*. Oxford: Pergamon.

Krashen, S.D. (1985). *The Input hypothesis: Issues & implications*. New York: Longman.

Lai, C., & Zhao, Y. (2005). The importance of input and the potential of technology for enhancing input. In Y. Zhao (Ed.) *Research in technology and second language learning. Developments and directions* (pp. 92-102). Greenwich: Information Age Publishing.

Laufer, B. (1997). The lexical plight in second language reading: Words you don't know, word you think you know and words you can't guess. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition: A rationale for pedagogy*. Cambridge: Cambridge University Press.

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

Laufer, B. & Nation, P. (1999). A vocabulary-size test controlled productive ability. *Language Testing*, 16, 33-51.

Laufer, B. & Hill, M. (2000). What lexical information do l2 learners select in a call dictionary & how does it affect word retention? *Language Learning & Technology, 3*(2), 58-76.

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

Laufer, B & Goldstein, Z. (2004). Testing vocabulary knowledge: Size, strength & computer adaptiveness. *Language Learning*, *54*(3), 399-436.

Leech, L., Barrette, K. C., & Morgan, G. (2008). *SPSS for introductory statistics: Use and interpretation*. New York: Lawrence Erlbaum Associate,

Lenders, O. G. (2008). Electronic glossing – is it worth the effort? *Computer Assisted Language Learning, 21*(5), pp. 457-481.

Levy, M. (1997). *Computer-assisted language learning: Context and conceptualisation*. Oxford: Oxford University Press.

Levy, M., & Stockwell, G. (2006). *CALL dimensions. Option & issues in computer-assisted language learning*. Mahwah: Lawrence Erlbaum Associates.

Li, J. (2010). Learning vocabulary via computer-assisted scaffolding for text processing. *Computer Assisted Language Learning*, *23*(3), 253-275.

Liaw M. L. (2009). Reading strategy awareness training to empower online reading. *The English Teacher*, XXXVIII, 145-162.

Lightbown, P. M., & Spada, N. (2006). *How languages are learned*. Oxford: Oxford University Press.

Lomicka, L.L. (1998). "To Gloss or not to gloss": An investigation of reading comprehension online. *Language Learning & Technology*, *1*(2), pp. 41-53.

Long, M. H. (1983). Linguistic & conversational adjustments to non-native speakers. *Studies in second language acquisition*, *5*(2), 177-193.

Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. Ritchie & T. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413-468). San Diego, CA: Academic Press.

Loschky, L. (1994). Comprehensible input & second language acquisition: What is the relationship? *Studies in Second Language Acquisition, 16,* 303-323.

Lyman-Hager, M & Davies, J. N. (1996). The case for computer-mediated reading: Une vie de boy. *The French Review, 69,* 775-790.

Mackey, A. (1999). Input, interaction, & second language development: An empirical study if question formation in ESL. *Studies in Second Language Acquisition,* 21, 557-587.

Mackey, A., & Gass, S. (2005). *Second Language Research*. New Jersey: Lawrence Erlbaum Associates Inc.

Malaysian Communications & Multimedia Commission. (2012). *Household use of the Internet survey 2011*. Retrieved from www. skmm.gov.my.

Martinez-lage, A. (1996) *Hypermedia technology for teaching reading in Technology-enhanced language learning*. Lincolnwood: National Textbook Company.

Mayer, R. (1997). Multi-media learning: Are we asking the right questions? *Educational Psychologists, 32*(1), 1-19.

Mayer, R. E. (2001). *Multimedia learning*. New York: Cambridge University Press.

Meara, P. (1992). Vocabulary in a second language. *Reading in a Foreign Language, 9,* 761-831.

Meara, P. (1997). Towards a new approach to modelling vocabulary acquisition. *Vocabulary: Description, acquisition and pedagogy*, 109-121.

Meara, P., & Fitzpatrick, T (2000). Lex30: An improved method for assessing productive vocabulary in an L2. *System, 28,* 19-30.

Melka Teichroew, F.J. (1982). Receptive versus productive vocabulary: A survey, *Interlanguage Studies Bulletin, 6,* 5-33.

Milton, J. (2009). *Measuring second language vocabulary acquisition*. Bristol: Multilingual Matters.

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

Mohamad Jafre Zainol Abidin, Majid Pour-Mohammad, Neda Sharbaf Shoar, See,T. H. C. & Afif Mohamad Jafre, (2011). A comparative study of using multimedia annotation and printed textual glossary in learning vocabulary. *International Journal of Learning & Development, 1*(1), 82-90.

Murphy, J. (1999). Surfers and Searchers: An examination of web-site visitors' clicking behaviour. *Cornell Hotel and Restaurant Administration Quarterly, 40*(2), 84-95. Retrieved from www.sciencedirect.com/science/journal/00108804.

Nagata, N. (1999). The effectiveness of computer-assisted interactive glosses. *Foreign Language Annals, 32*(4), 369-479.

Nation, P. (1990). *Teaching and learning vocabulary*. Rowley, MA: Newbury House.

Nation, P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.

O'Donnell, M. E. (2013). Second language learners' use of marginal glosses. *Foreign Language Annals, 45*(4), 543-563.

Oh, J., Robinson, H. R., & Lee, J.Y. (2013). Page flipping vs. clicking: The impact of naturally mapped interaction technique on user learning and attitudes. *Computers in Human Behaviour, 29*(4), 1334-1341. Retrieved from www.eslsevier.com/locate /comphumbeh.

Ortega, L. (2009). *Understanding second language acquisition.* London: Hodder Education.

Palmberg, R. (1987). Patterns of vocabulary development in foreign language learners. *Studies in Second Language Acquisition, 9*, 201-220.

Park, E. S. (2002). On three potential sources of comprehensible input for second language acquisition. *Working papers in TESOL & Applied Linguistics, 2*(3), Retrieved from http://journals.tc-library.org/index.php/tesol/issue.

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

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

Plass, J. L., Chun, D. M., Mayer, R. E., & Leutner, D (1998). Supporting visual and verbal learning preferences in a second language multimedia learning environment. *Journal of Educational Psychology, 90*(1), 25-36.

Plass, J. L., & Jones, L. C. (2005). Multimedia learning in second language acquisition. In R.E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning*. Cambridge: Cambridge University Press.

Pica, T. (1994). Research on negotiation: What does it reveal about second-language learning conditions, processes, & outcomes?" *Language Learning,* 44, 493-527.

Pica, T. (1996). *Second language learning through interaction: Multiple Perspectives*. ERIC Clearinghouse.

Pica, T. (2005). Second language acquisition research and applied linguistics. *Handbook of Research in Second Language Teaching and Learning*, 263-280.

Pignot-Shahov, V. (2012). Measuring L2 Receptive and Productive Vocabulary Knowledge. *LANGUAGE*, *4*, 37-45.

Pimsleur, P. (1967). A Memory schedule. *The Modern Language Journal*, *51*(2), 73-75.

Read, J. (2000). *Assessing vocabulary*. Cambridge: Cambridge University Press.

Read, J. (2004). Research in teaching vocabulary. *Annual Review of Applied Linguistics, 24*(1), 146-161.

Reinking, D., & Leu, D. J. (1996). Bringing insights from reading research to research on electronic learning environments. *Advances in Discourse Processes, 58*, 43-76.

Robb, W. (1999). "What's in a gloss?" *Language Learning & Technology*, *2*(2), 94-101. Retrieved from http://llt.msu.edu/vol2num2/roby/index/htmal.

Rudestam, K. E., & Newton, R. R. (2001). *Surviving your dissertation: a comprehensive guide to content and process*. Thousand Oaks, California: Sage Publications.

Saville-Troike, M. (2006). *Introducing second language acquisition.* New York: Cambridge University Press.

Shehadeh, A. (2002). Comprehensible output from occurrence to acquisition: An agenda for acquisitional research. *Language Learning, 52*, 597-647.

Schmitt, N. (1998). Tracking the incidental acquisition of second language vocabulary: A longitudinal study. *Language Learning, 48*(2), 281-317.

Schmitt, N. (2000). *Vocabulary in language teaching*. Cambridge: Cambridge University Press.

Schmitt, N. (Ed.). (2002). An *introduction to applied linguistics*. London: Arnold.

Schmitt, N., (2010). *Researching vocabulary. A vocabulary research manual*. Hampshire: Palgrave Macmillan

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

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

Schmidt, R. (1995). Consciousness and foreign language learning: A tutorial on the role of attention and awareness in learning. *Attention and awareness in foreign language learning*, 1-63.

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

Shahrokni, S. A. (2009). Second language incidental vocabulary learning: The effect of online textual, pictorial, and textual pictorial glosses. *TESLEJ*, *13*(3). Retrieved from http://www.tesl-ej.org/wordpress/

Smith, S. (1986). Comprehension vs. acquisition: Two ways of processing input. *Applied Linguistics, 7,* 238-56.

Smith, S. (1993). Input enhancement in instructed SLA: Theoretical bases. *Studies in Second Language Acquisition*, *15*(2), 165-179.

Smith, B. (2004). Computer-mediated negotiated interaction and lexical acquisition. *Studies in Second Language Acquisition*, *26*(3), 365-398.

Stoehr, L. (1999). *The effects of built-in comprehension aids in a CALL program on student readers' understanding of a foreign language literary text*. (Unpublished PhD thesis). University of Texas.

Swain, M. (1985). Communicative competence: some roles of comprehensible input & comprehensible output in its development. In S.M. Gass & C.G. Madden (Eds.), *Input in second language acquisition* (pp. 235-253). Rowley, MA: Newbury House.

Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidhofer (Eds.), *For H.G. Widdowson: Principles and practice in the study of language* (pp. 125-144). Oxford: Oxford University Press.

Swain, M. (2005). The output hypothesis: Theory and research. In E. Hinkel (Ed.), *Handbook on research in second language learning and teaching.* Mahwah, NJ: Lawrence Erlbaum Associates.

Taylor, A. (2006). The effects of CALL versus traditional L1 glosses on L2 reading comprehension. *CALICO Journal*, *23*(2), 309-318.

Taylor, A. M. (2010). Glossing is sometimes a distraction: Comments on Cheng & Good (2009*). Reading in a Foreign Language*, *2* (22), 353-354.

Tomlin, R. S., & Villa, V. (1994). Attention in cognitive science and second language acquisition. *Studies in Second Language Acquisition*, *16*(2), 183-203.

Felix, U. (2008). The unreasonable effectiveness of CALL: What have we learned in two decades of research. *ReCALL*, 20, 141-161 doi: 1017/S0958344008000323

Van Patten, B., & Williams, J. (2007). Theories *in second language acquisition*. New Jersey: Lawrence Erlbaum Associates Publishers.

VanPatten, B. (2012). Input processing. In S.M. Gass & A. Mackey (Eds.). *The Routledge handbook of second language acquisition*. (pp. 268-281). New York: Routledge.

Warschauer, M. (1999). Electronic *literacies. Language, culture & power in online education.* Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

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

Wesche, M. B. (1994). Input & interaction in language acquisition. In C. Gallaway & B. J. Richards (Eds.), *Input and interaction in language acquisition*. (pp. 238-249). Cambridge: Cambridge University Press.

White, L. (1987). Against comprehensible input: the input hypothesis and the development of second language competence. *Applied Linguistics*, *8*, 95-110.

Yanguas, I. (2009). Multimedia glosses and their effect on L2 text comprehension and vocabulary learning. *Language Learning & Technology*, *13*(2), 48-67.

Yeh, Y., & Wang, C.W. (2003). Effects of multimedia vocabulary annotations and learning styles and learning vocabulary. *CALICO* Journal, *21*(1), 131-141. Available at https://calico.org/journalTOC.php

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

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

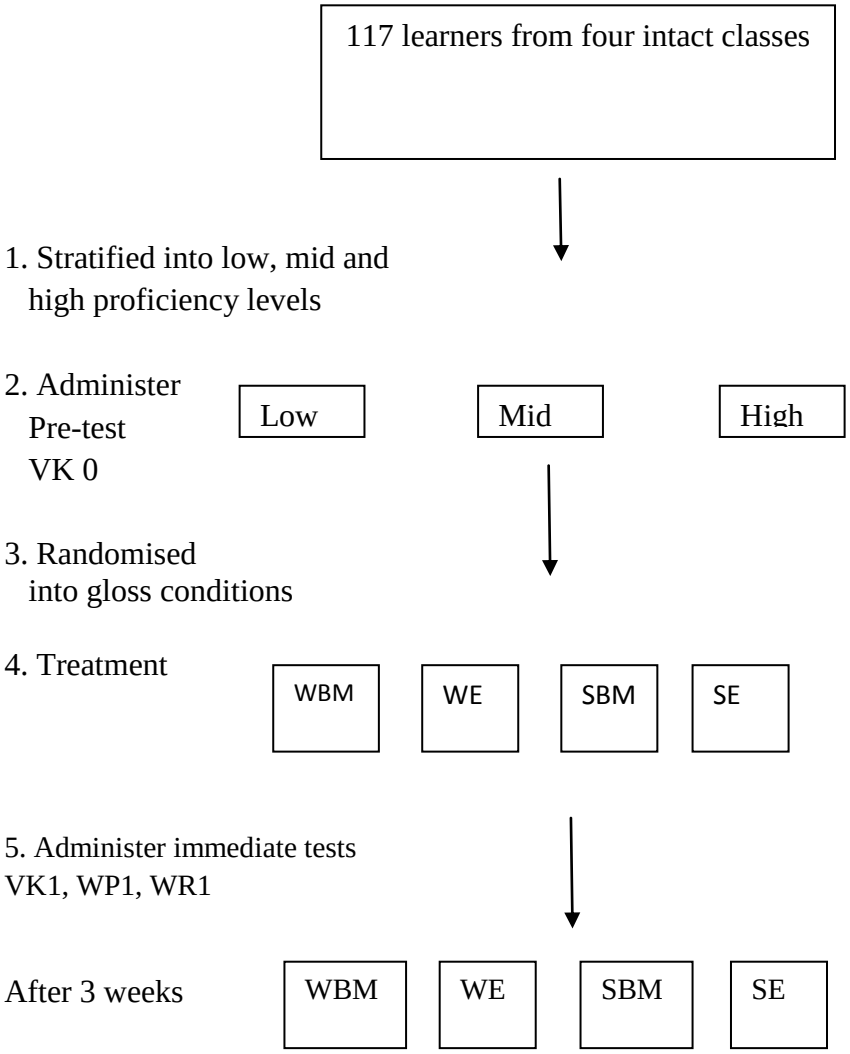
Youngs, B., Ducate, L., & Arnold, N. (2011). Linking second language acquisition, CALL, and language pedagogy. In N. Arnold and L. Ducate (Eds.), *Present and future promises of CALL: From theory and research to new directions in language teaching* (pp. 23-60). Texas: CALICO.

Yun, J. (2011). The effects of hypertext glosses on L2 vocabulary acquisition: a meta analysis. *Computer Assisted Language Learning*, *1* (24), 39-58.

Zhao, Y. (Ed.). (2005). *Research in technology and second language education: Developments and directions*. Greenwich: Information Age Publishing.

APPENDIX A

Research Procedure



Key

VK1 perceived vocbulary knowledge test 1

VK2 perceived vocbulary knowledge test 2

VK3 perceived vocbulary knowledge test 3

WP1 word production test1

WP2 word production test2

WR1 word recognition test 1

WR2 word recognition test 2

APPENDIX B

What is involved in knowing a word?

Form	spoken	R	What does the word sound like?
		P	How is the word pronounced
	written	R	What does the word look like?
		P	How is the word written and spelled?
	Word parts	R	What parts are recognizable in this word?
		P	What word parts are needed to express the meaning?
Meaning	form and meaning	R	What meaning does this word form signal?
		P	What word form can be used to express this meaning?
	concept and referents	R	What is included in the concept?
		P	What items can the concept refer to?
	associations	R	What other words does this make us think of?
		P	What other words could we use instead of this one?
Use	grammatical functions	P	In what patterns does the word occur?
		R	In what patterns must we use this word?
	collocations	R	What words or types of words occur with this one?
		P	What words or types of words must we use with this one?

	Constraints on use	R	Where, when, and how often would we expect to meet this word?
		P	Where, when, and how often can we use this word?

R stands for receptive knowledge, P stands for productive knowledge

Taken from Nation, 2001, p.26.

APPENDIX C

The Text – A Scary Night

It is a cold rainy night. It is midnight, and it is very quiet. I am still awake and studying. I have a test tomorrow. I need to read another two chapters. I finished one chapter a few minutes ago and I will start on the next chapter. It is difficult. I cannot pass the test. What do I do? Shall I keep studying? Can I take the test some other time? Shall I give up? I am **pondering** many things. Life as a student is indeed difficult. I think my head is going to **burst**. But I must read on.

Suddenly, some noise **startles** me. I stopped reading and looked around me. Something **shattered** on the ground outside my house. I look at the window. Wait! What is that? I see a light across the street. It is from a new house across the road. This is strange. Mr. and Mrs. Leong are on holiday now. They asked me to **rake** the garden for them while they are gone. Nobody should be there. Oh, I see the light again.

Then, I try to concentrate on my reading. But I can't. Suddenly, it **dawns** on me. There must be someone in the house. Who could that be? What a fool I am? Of course, someone is **burglarizing** the house! I am afraid. What should I do now? I have to call the police. I **dash** to the phone and call the police.

After ten minutes, the police arrive. They enter the house. As the police search the house, someone hides outside the house. The police yell, "Stop, right there!" But the man with a black mask runs into the jungle behind the house. Then, he **tumbles** down the hill. The police run after him and catch him. They take off the mask. The burglar **grins** first, surprise at the bright lights of flash lights of the police. Then, he starts to **sob**.

Two policemen come to my house. The first one looks very serious. He does not greet me. He just asks for my name. He says, "Thank you for calling us about this problem." The other one is friendlier. He inquires about a few things. He wants to know when I first saw the light. He **scribbles** some notes.

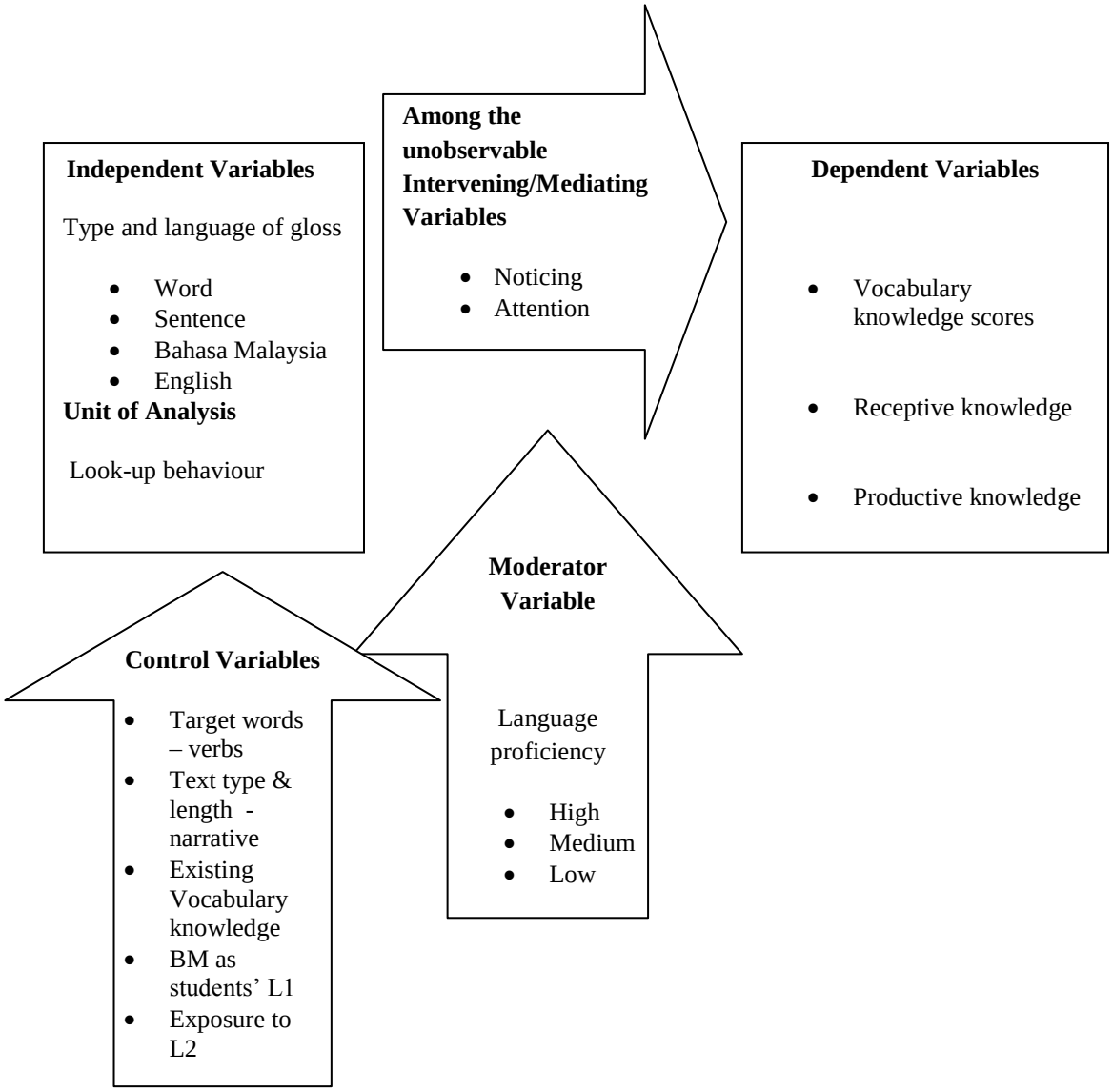
The policemen are gone, and everything is quiet now. What a strange night! I am glad this is over, but I am still **s**hivering a little. I go into the kitchen and pour some hot coffee into a cup. This might help me. I can't study any longer and can't sleep right away too.

I decide to read a book. I bought it at a bookstore yesterday. The title is "American Short Stories." I look at the first story and I **gape** at the title. It is "My Life as a Burglar" by A Man with a Black Mask.

-end-

APPENDIX D

Variables for the Study



APPENDIX E

The Target words and their Frequency Bands

Target Words	**Frequency Banding**
Ponder	BNC - 8,000
Burst	BNC - 3,000
Startle	BNC - 6,000
Shatter	BNC - 4,000
Rake	Off list
Burglarize	BNC - 3,000
Dash	BNC - 3,000
Tumble	BNC - 3,000
Grin	BNC - 4,000
Sob	BNC - 7,000
Scribble	BNC - 3,000
Quiver	BNC - 9,000
Gape	BNC - 9,000

APPENDIX F

Robb's Taxonomy

I. Gloss authorship

1. Learners
2. Professionals
3. Instructors
4. Materials developer

II. Gloss presentation
1. Priming
2. Prompting

III. Gloss functions
1. Procedural
   (a) Metacognitive
   (b) Highlighting
   (c) Clarifying
2. Declarative
   (a) Encyclopaedic
   (b) Linguistic
      (i) Lexical
         - signification
         -value
      (ii) Syntactical

III. Gloss focus
1. Textual
2. Extra textual

IV. Gloss language
1. L1
2. L2
3. L3

VI. Gloss form

1. Verbal
2. Visual
   (a) Image
   (b) Icon
   (c) Video
      - with sound
      - without sound
      - Audio only

APPENDIX G

Perceived Vocabulary Knowledge Test

Name: _____ UiTM No. _____

Your SPM English Grade: _____ Group: ACS113 _____

For the following words, tick √ the most suitable category that best suits you as provided by the scale below.

0. I **definitely don't know** what this words means
1. I **am not really sure** what this word means
2. I **think I know** what this word means
3. I **definitely know** what this word means

Bagi perkataan-perkataan di bawah, sila tandakan √ pada kategori yang paling sesuai dengan anda berdasarkan skala di bawah.

*0. Saya **amat pasti saya tidak tahu** makna perkataan ini*
*1. Saya **tidak berapa pasti** makna perkataan ini*
*2. Saya **fikir saya tahu** makna perkataan ini*
*3. Saya **amat pasti saya tahu** makna perkataan ini*

No.	Word	0	1	2	3
1.	Tumble				
2.	Grin				
3.	Sob				
4.	Scribble				
5.	Quiver				
6.	Gape				
7.	Dive				
8.	Swallog				

9.	Energise				
10.	Build				
11.	Wonder				
12.	Rake				
13.	Manitor				
14.	Ponder				
15.	Burst				
16.	Startle				
17.	Shatter				
18.	Smere				
19.	Burglarize				
20.	Dash				

Bagi perkataan-perkataan di atas, sila tandakan √ pada kategori yang paling sesuai dengan anda berdasarkan skala di bawah.

*0. Saya **amat pasti saya tidak tahu** makna perkataan ini*

*1. Saya **tidak berapa pasti** makna perkataan ini*

*2. Saya **fikir saya tahu** makna perkataan ini*

*3. Saya **amat pasti saya tahu** makna perkataan ini*

APPENDIX H

Productive Vocabulary Knowledge test

Name:_____ Group: AAP _____

UiTM No:_____ SPM EnglishGrade:_____

Fill in the blanks of the sentences with suitable words from the story "A Scary Night" which you have read before. The initial letters are given to help you choose the correct word.

1. As they had no shelter, the victims of the flood **qu**_____ in the cold.
2. Siti is a serious girl and often **po**_____ on the things that happen in her life.
3. In life, it is usual to **tu**_____ over obstacles.
4. There was too much water in the dam after the heavy rain that the dam **bu**_____.
5. After the accident, pieces of **sh**_____ glass was all over the road.
6. The baby was **st**_____ by the thunder.
7. He was pleased he had won the competition and had a wide **gr**_____ on his face.
8. When his name was announced as the winner he was surprised and only managed to **ga**_____ at the audience.
9. It is common to find many safety features in home today to avoid being **bu**_____.
10. It was raining heavily but Ali had to **da**_____ to the pharmacy before it closed.
11. He had to **ra**_____ his neighbours' garden for some pocket money during the holidays.
12. Ahmad felt sorry for her after she started to **so**_____.
13. Doctors are often thought to **sc** _____ as most people cannot read their hand writing.

APPENDIX I

Receptive Vocabulary Knowledge test

Name:_____ Group: AAP
UiTM No:_____ SPM English Grade:

Directions: Please match the words below with the correct meaning. Put [X] in the box.

1.Burglarize

[ ] To move or act very fast; to hurry

[ ] To clean something very quickly

[ ] To enter a building and steal something

[ ] To say something loudly and angrily

2.Sob

[ ] To smile widely

[ ] To feel very sorry

[ ] To cry without control

[ ] To run away quickly

3.Gape

[ ] To open the mouth widely

[ ] To talk angrily

[ ] To laugh loudly

[ ] To shake uncontrollably

4.Burst

[ ] To break open suddenly

[ ] To shut suddenly

[ ] To carry many heavy things

[ ] To think of problems

5.Rake

[ ] To cut grass using a machine

[ ] To gather leaves with a special tool

[ ] To trim garden trees and hedges

[ ] To water the plants

6.Shatter

[ ] To break into small pieces

[ ] To glue into small pieces

[ ] To fix from small pieces

[ ] To hit into small pieces

7.Tumble

[ ] To say something softly

[ ] To hide outside silently

[ ] To run away quickly

[ ] To fall quickly

8.Dash

[ ] To move with speed

[ ] To call someone loudly

[ ] To talk with someone angrily

[ ] To shout loudly

9.Startle

[ ] To break something into pieces

[ ] To make a loud noise

[ ] To become very quiet

[ ] To surprise or scare someone

10.Scribble

[ ] To write quickly

[ ] To look at something carefully

[ ] To write slowly

[ ] To read carelessly

11.Grin

[ ] To cry softly

[ ] To talk softly

[ ] To stare at someone

[ ] To have a big smile

12.Ponder

[ ] To study hard

[ ] To think carefully

[ ] To read quickly

[ ] To open suddenly

13.Quiver

[ ] To shiver

[ ] To laugh

[ ] To worry

[ ] To feel

APPENDIX J

Consent Form

Borang Persetujuan

Assalamualaikum dan selamat sejahtera pelajar-pelajar

Borang Persetujuan ini adalah untuk meminta bantuan dan kerjasama daripada pelajar-pelajar untuk menyertai penyelidikan yang saya sedang jalankan. Kajian ini adalah untuk tesis doktor falsafah saya di Universiti Malaya yang berjudul: *Learner-computer textual gloss interactions for second language acquisition.* Penyelidikan ini penting untuk meluaskan pengetahuan mengenai penggunaan computer dalam pembelajaran perbendahraan kata Bahasa Inggeris.

Sebarang maklutmat yang pelajar-pelajar beri adalah untuk kajian ini sahaja. Maklumat yang diperolehi adalah sulit dan ia tidak akan menjejaskan prestasi mahupun keputusan matapelajaran Bahasa Inggeris yang pelajar-pelajar sedang ikuti.

Saya berharap pelajar-pelajar akan setuju untuk turut serta dalam kajian ini.

Sekian. Terima kasih.

Mohamad Ali Yusuf

☐ Saya setuju untuk menyertai penyelidikan ini dan maklumat yang diperolehi digunakan oleh Mohamad Ali Yusuf untuk tesisnya dan lain-lain penerbitan.

☐ Saya tidak setuju untuk menyertai penyelidikan ini.

(Sila tandakan √ pada kotak yang bersesuaian)

Nama : _____

Tarikh: _____

Tandatangan: _____