

CHAPTER 1: INTRODUCTION

1.1 Background to the Study

1.1.1 L2 Vocabulary Research

For the past decades, more information has become available on L2 vocabulary research (Meara, 1980; Maiguashca, 1993; Laufer, 1998). However, from the 1960s to 1980s, L2 research especially on vocabulary has been undervalued (Maiguashca, 1993). This event prompted scholars to raise concern over L2 vocabulary research and refer this linguistic aspect as being overlooked and having no relational cause to language learning (Maiguascha, 1993). Furthermore, vocabulary development was regarded as a redundant learning process in the classroom as language learners were assumed to acquire it subconsciously. Such perspectives may be caused by two plausible reasons: (1) grammatical structures and phonology established themselves in the linguistic arena earlier than semantics, and (2) there is no systematic structures found in vocabulary in comparison to grammar and phonology. In other words, learning vocabulary is not as specific and clear as learning grammatical features of the language (Maiguascha, 1993).

Nevertheless, before L2 vocabulary acquisition research gained popularity, researches were essentially focused on frequency counts and mnemonic techniques. The 1950s saw large volumes of published studies in the investigation of vocabulary size as demonstrated by the General Service List (West, 1953) in determining 2,000 high frequency words extracted from a corpus of written English. However, the purpose of the aforementioned study was not driven by any theoretical frameworks but rather by learning materials involved during the language learning process. By ascertaining the high and low frequency words, it may assist language material writers and editors in determining which words should be included in the materials. On the other hand, mnemonic techniques are categorized as one of the strategies in learning vocabulary and gained attention in the 1970s. Mnemonic technique has been regarded by Nation (1990)

to challenge the frequency count assumption that language learners could learn a gargantuan amount of lexical items (Meara, 1980). Meara's (1980) study was one of the pivotal studies that marked the emerging key word method. Despite introducing mnemonic techniques, Meara (1980) asserts that frequency count and mnemonic techniques merely deal with the minor aspects of the vocabulary and not the principal areas, which is the vocabulary learning process.

Over the past decade, L2 vocabulary research extended its research scope to psycholinguistic factors that promotes L2 vocabulary learning and acquisition (Ellis & Beaton, 1993; Laufer, 1998), methods in measuring vocabulary knowledge (Wesche & Paribakht, 1996; Read, 1997), receptive and productive vocabulary acquisition (Melka Teichroew, 1982) and many others. Since then, vocabulary research has established itself as part of an important language learning aspect in the linguistic arena (Manguashca, 1993; Laufer, 1998).

1.1.2 L2 Vocabulary Acquisition

Many researchers have argued that vocabulary acquisition is one of the fundamental aspects in the field of SLA (Coady, 1993; Nation, 1982, 2001; Meara, 1980). Beck and Mckeown (1991) demonstrated that L2 vocabulary acquisition enables access to understand texts and communication. Apart from that, numerous reports have shown that one of the communication barriers faced by L2 learners are caused by the choice of lexical items (Swain & Lapkin, 1995; Meara, 1980; Nation, 1982; Mackey, Gass, & McDonough, 2000).

While vocabulary knowledge remains vital for L2 communication, researchers and educators alike have yet to reach a consensus in deciding the best method to learn and acquire L2 vocabulary. To illustrate the divide in relation to second language acquisition (SLA) theories, one of the most important contributions in the 1980s was the different benefits of input and output in L2 learning (Krashen, 1982; Long, 1983;

Swain, 1985). Prior to that, Wagner-Gough and Hatch (1975) suggested that input is the only independent variable that promotes language acquisition, which initially caused a stir in the linguistic arena (Krashen, 1982; Long, 1983). The 1970s witnessed input as active determining element while output, which primarily deals with practice after the learning and acquisition process, assumes a minor role in SLA (Krashen, 1982; Gass & Varonis, 1994). However in 1985, Swain postulated that output might play an equal role as input in L2 acquisition. In other words, the input hypothesis (Krashen, 1982; Long, 1983) insinuates that in order for second language acquisition to take place, input must be made comprehensible to the learners while the output hypothesis as proposed by Swain (1985) argued that learners should be provided ample room to produce linguistic forms in the L2 to trigger the acquisition process.

Research to date, however, has not offered conclusive findings as to which independent variable; input, output or the combination of both, offers the best approach for L2 acquisition. Therefore, there has been an increasing amount of literature in determining the effectiveness of different oral interactions on L2 acquisition (Pica, 1994; Ellis, Tanaka & Yamazaki, 1994; Van de Branden, 1997; de la Fuente, 2002; Loschky, 1994; Oliver, 1998; Ellis & He, 1999; Gass & Torres, 2005; Mackey et al., 2010; Ng & Sheila Marie Sappathy, 2010; Shintani, 2011). The researches involved the manipulation of input and output conditions and were predominantly focusing on pre-modified input, negotiated input and also output production. In other words, researches claimed that negotiated input provides better comprehension than pre-modified input but not for vocabulary acquisition (Loschky, 1994; Ellis, 1994; Ellis, Tanaka, & Yamazaki, 1994). Additionally, output production is more beneficial in relation to L2 comprehension, receptive acquisition and productive acquisition than pre-modified input and negotiated input with no statistical difference between the aforementioned two input conditions. Finally, although the researches have clearly shown that output

production and negotiated input gained greater advantage in terms of L2 comprehension, it was remained undecided for receptive and productive acquisition (Ellis, 1985; Loschky, 1994; Ellis & He, 1999), which warrants further scrutiny into the three oral interactions in this present study.

1.2 Problem Statement

In this age of globalization, there has been an increase in the usage of English throughout the world (Platt & Weber, 1980). Without a doubt, the English language plays a pertinent role in the progress and growth in major developmental areas in Malaysia (Rahmah Hashim, 1996). The relative importance of English in Malaysia contributes to certain domains such as the business sector, employment sections, education, politics, tourism, law, media and many others.

From the education perspective, the government has been consistently emphasizing on the importance of English language by incorporating English as a second language from primary to tertiary level as it is highly important in terms of nation building and global competitiveness. With the current implementation of KSSR, the Standard Curriculum for Primary School, there has been an increase in the time for English lessons from the previous Integrated Curriculum for Primary School, known as KBSR. However, despite formal learning of English for six consecutive years, Malaysian pupils are unable to attain reasonable communicative competence while transitioning from primary school to secondary school (Nor Hashimah Jalaludin, Norsimah Mat Awal & Kesumawati Abu Bakar, 2008; Saadiyah Darus & Kaladevi Subramaniam, 2009; Pak, 2012; Ramiza Darmi & Albion, 2013). While delving into the English language learning issues in Malaysian ESL classrooms, the aforementioned reports saw that pupils struggled in understanding and communicating with their English teachers and their fellow peers who possessed higher level of English proficiency. This issue snowball into a detrimental issue at a macro level. This resulted

in potential employers rating graduates in Malaysia as unemployable due to poor performance in communication skills (World Bank, 2009; The Borneo Post, 2012).

Among the many reasons for the unsatisfactory communication skills is the lack of emphasis on communicative language learning activities for ESL learners in primary national schools (Saadiyah Darus, 2010; Ng, 2013). English lessons that incorporate communicative activities were deemed unfavourable as errors were ignored and fluency stressed over accuracy. On the other hand, form-based tasks and knowledge-based examinations were given importance to gear pupils up for the public examination, widely known as UPSR (Ujian Pencapaian Sekolah Rendah) (Reza Raissi & Fazirah Mohd Nor; 2013). Overemphasis on certain aspects of the language such as focusing on the grammatical structures and writing may subconsciously cause teachers to neglect on other areas like listening, speaking and L2 vocabulary development. In line with Nunan's (1991) view, language use opportunities and successful communication are dependent upon the mastery of L2 vocabulary. Therefore, pupils should learn and acquire lists of vocabulary that would be deemed sufficient to engage in communication and understanding a message (Nunan, 1991). According to Hatch's (1978) work, the communicative process of negotiation promotes second language comprehension and the type of task that is normally involved emulates the information gap format to push learners to communicate in classrooms (Doughty & Pica, 1986). Negotiated input has also been suggested to promote L2 vocabulary acquisition in terms of retention whereby language learners will have to ability to hold the vocabulary for short-term and long-term retrieval in their memories (Ng & Sheila Marie Sappathy, 2010) with particular reference to nouns.

The inquiries that motivated the current study are based on de la Fuente's study (2002), which examined the effects of three different oral interactions, which are pre-modified input, negotiated input and output production on L2 vocabulary

comprehension and acquisition. Findings from the study revealed that negotiated input promoted better comprehension and output production produced better acquisition results. Aside from exploring L2 vocabulary comprehension, this study is worth exploring because it deals specifically on two types of acquisition; receptive and productive acquisition. The first underlying theoretical framework of the research questions and hypotheses is that for second language acquisition to take place, comprehensible input is required based on the input hypothesis (Krashen, 1985) and interaction hypothesis (Long, 1983). The second underlying proposition is that pupils should be given the opportunity to produce linguistic forms in the L2 to promote L2 acquisition as postulated by Swain (1985) in her output hypothesis.

In this regard, there is a need to investigate the role of pre-modified input, negotiated interaction and output production in L2 vocabulary comprehension and acquisition. It was not uncommon that previous studies on negotiation essentially comprised of young adults with multiple L1 backgrounds (Ellis, Tanaka & Yamazaki, 1994; Loschky, 1994; Ellis and He, 1999; Mackey, 2010; de la Fuente, 2002). Thus, the current study attempts to investigate L2 vocabulary comprehension and acquisition with respect to children (primary ESL learners) who are of similar ethnic backgrounds and mother tongue. The participants are subjected to a two-way information gap format task in order to provide quantitative data as evidentiary conjecture to measure L2 vocabulary comprehension and acquisition while negotiating. L2 vocabulary acquisition may well be the one of the steps in enhancing communicative competence among primary ESL learners in Malaysia.

1.3 Purpose of the Study

The purpose of this study is to investigate the effects of different oral interactions on new L2 vocabulary acquisition among Primary ESL learners in Kuala Langat, Malaysia. The objectives of the study are: -

- i) to determine the effectiveness of negotiated input in vocabulary comprehension.
- ii) to determine the effectiveness of negotiated input with and without output in receptive vocabulary acquisition.
- iii) to determine the effectiveness of negotiated input with and without output in productive vocabulary acquisition.

1.4 Research Questions and Hypotheses

To substantiate the findings of the study, three research questions with their corresponding hypotheses were formed and tested. The hypotheses and research questions are as follows:

a) Vocabulary Comprehension

Learners who are exposed to pre-modified input with negotiated interaction will attain higher level of vocabulary comprehension than those learners who are exposed to pre-modified input only.

1. What are the relative effects of pre-modified input and negotiated interaction on L2 vocabulary comprehension?

b) Receptive Vocabulary Acquisition

Learners who are exposed to pre-modified input with negotiated interaction with or without output production will attain higher level of receptive vocabulary acquisition than those learners who are exposed to pre-modified input only.

2. i) What are the relative effects of pre-modified input, negotiated interaction with and without output production on L2 receptive vocabulary acquisition?
- ii) If there is an effect, how does it affect L2 vocabulary acquisition in terms of retention (time)?

c) Productive Vocabulary Acquisition

Learners who are exposed to pre-modified input with negotiated interaction with or without output production will attain higher level of productive vocabulary acquisition than those learners who are exposed to pre-modified input only.

3. i) What are the relative effects of pre-modified input, negotiated interaction with and without output production on L2 productive vocabulary acquisition?
- ii) If there is an effect, how does it affect L2 vocabulary acquisition in terms of retention (time)?

1.5 Significance of the Study

This study is significant in three ways. The findings from this research provide insights not only to language researchers but also to language teachers and the school authority. Firstly, it is hoped that the study may contribute to the limited literature of negotiated interaction and output production especially for young children and the primary ESL learning context. The findings of the study may support its effectiveness towards L2 vocabulary acquisition, in association to nouns, among primary ESL learners in Malaysia. This present study may help to pave ways for future researches to determine the effectiveness of input and output specifically in L2 vocabulary acquisition among young language learners. Apart from that, the intricacy involved in dealing with young ESL children in relation to the choice of lexical items, the types of interactive tasks, the procedures and the outlay of each tasks are provided to aid future researches in similar endeavors.

Secondly, the present study may also assist language learners by raising awareness on the benefits of communicative activities to language teachers. As this experiment was to determine the effectiveness of three input conditions; pre-modified input, negotiated interaction and output production, the findings will provide language

teachers insights on specific methods to promote L2 vocabulary acquisition in order to enhance the pupils' communicative skills in the classroom settings. As iterated by Ellis (1997), experiments on different input conditions could provide a sense of familiarity by assisting language teachers in identifying the strengths and weaknesses on the different input conditions. Language teachers may conduct activities that incorporate input, negotiated interaction and output to promote L2 vocabulary acquisition.

Thirdly, the findings of this study may inform the school authority and the general public on the importance of incorporating negotiated interaction and output production as part of the syllabus design in English lessons. Primary ESL classrooms in Malaysia place emphasis on form-based tasks and knowledge-based examinations. This is primarily driven by the public pressure to perform in UPSR (The Sun Daily, 2014), a public examination for pupils who are attending primary schools in Malaysia. It is hoped that the findings of this study will help the school authorities to become more attuned and sensitive to the needs of the pupils in the communication aspect. This eases the teachers' dilemma in selecting the best approach to learn in primary ESL classrooms. Aside from focusing on form-based tasks, communicative activities may be regarded as a commodity to improve the speaking skills of young ESL learners.

1.6 Scope and Limitations

Due to time and financial constraints, the proposed study will be limited to 45 pupils from primary national type schools in Kuala Langat, Malaysia. They are currently in Primary 3. Apart from that, the target items tested will be focusing on a specific word class, a list consisting of 10 concrete nouns in the English Language. The selection of the nouns, which includes 10 types of vegetables, is based on the topical word lists to be learnt as stipulated in the Standard Document (KSSR, 2011). The study will investigate the acquisition of the singular form and basic meaning of the nouns (Nation, 2001). (Therefore, the findings of the study may not be generalizable to other

word classes (eg. verbs, adjectives, adverbs) and other aspects of lexical acquisition (eg. associations, collocations).

1.7 Operational Definitions

The following independent and dependent variables involved in the present study are provided with brief definitions. Chapter 2 presents detailed descriptions of each key term that is relevant to the purpose of the study.

1.7.1 Pre-modified Input

Pre-modified input is described as input that is modified by the instructor to suit the target learners' current level of comprehension. This minimizes comprehension barrier (Pica et. al., 1987; Pica, 1994).

1.7.2 Negotiation

Negotiation is defined as an interactional modification to achieve input or message comprehensibility (Krashen, 1980; 1985) between interlocutors and learners. Negotiation may be in forms of repetition, paraphrasing, form adjustments, syntactical modification and substitution of words (Pica, 1994)

1.7.3 Output production

Output production refers to how learners are provided the opportunity to produce in the target language (the English Language) for language use opportunities such as speaking or writing (Swain, 1985; Long, 1996).

1.7.4 Comprehension

Comprehension may not necessarily lead to acquisition (Sharwood Smith, 1986). Comprehension requires learners to decode a particular message that has been linguistically encoded (Sharwood Smith, 1986; White, 1987; Ellis, 1991).

1.7.5 Acquisition

Acquisition refers to the capacity of the learner to create new mental structures of new linguistics forms (Gass & Varonis, 1994; Schmidt & Frota, 1986; Schmidt, 1990). For

this study, the term vocabulary acquisition will be defined closely based on Nation's (2001) word knowledge, which includes knowledge of form and knowledge of word meaning. Knowledge of form encompasses knowing what word looks like and what it sounds like in its phonological form. Knowledge of word meaning is divided into three parts. However, the term vocabulary acquisition will be closely linked to the first subdivision, which is form and meaning. It encompasses the ability to link the form to a meaning by forming a link between the English Language to its translation the learners' native language (Malay) (Milton, 2009).

1.7.6 Retention

Acquisition is often linked to retention, where language learners have the ability to hold the target words for short-term and long-term retrieval.

1.7.7 Receptive Vocabulary Acquisition

In a language-learning context, receptive acquisition is linked to the ability to understand a word while listening and reading (passive knowledge) (Nation, 2001; Laufer & Goldstein, 2004). The current study focuses on the receptive vocabulary acquisition of knowledge form (Nation, 2001; Milton, 2009).

1.7.8 Productive Vocabulary Acquisition

Productive acquisition is linked to the ability to use a word in speaking and writing (active knowledge) (Nation, 2001; Laufer & Goldstein, 2004). The current study focuses on the productive vocabulary acquisition of knowledge meaning (Nation, 2001; Milton, 2009).

1.8 Conclusion

Chapter 2 will provide a thorough review of literature at a macro and micro level and is divided into three distinctive parts. The first part covers the historical background on the input conditions investigated in L2 vocabulary acquisition. The second part consists of the theoretical frameworks that this present study is grounded on, which are

the three language learning theories in the field of SLA. The final portion of Chapter 2 encompasses the relevant studies which warrants further research on the three independent variables; pre-modified input, negotiated input and output production in L2 vocabulary acquisition among primary ESL learners in Malaysia. Chapter 3 will outline the detailed and systematic information of the research design, participants, testing instruments, step-by-step procedure of the treatment and measurement process and a comprehensive description of conducting data analysis. Chapter 4 will present the results of the tests scores through descriptive and inferential statistics. The report will show how each research question and hypothesis are being answered and analyzed. Finally, Chapter 5 attempts to provide a summary, interpretation in the light of relevant literature, discussion relevant to the results of the study. Apart from that, the implications to primary ESL learners and suggestions for further research are presented to conclude the chapter.

CHAPTER 2: LITERATURE REVIEW

2.1 Introduction

This chapter is divided into three main sections. The first section will provide a historical account of negotiated interaction and L2 vocabulary acquisition via relevant literature reviews. This is followed by the various theories used extensively to frame this study, namely the input hypothesis, the interaction hypothesis and the output hypothesis. Finally, section three will identify relevant studies in relation to negotiation and L2 vocabulary development specifically outlining the various methodologies used and the outcomes.

2.2 Historical Background

2.2.1 Negotiation and Second Language Acquisition

In line with the interaction hypothesis (Long, 1981; 1985), the underlying premise of Hatch's (1978) work is that the communicative process of negotiation promotes second language acquisition (SLA). Her seminal papers (Hatch, 1978) predominantly analyze the discourse between caretakers and language learners and how constant interaction assists the acquiring of new linguistic forms (Breen & Littlejohn, 2000). Hatch's (1978; 1983) formative idea was driven by a different postulation that L2 linguistic forms may developed from interaction as opposed to how learning L2 linguistic forms may aid in the process of L2 communication.

Hatch's work (1978) was deemed pivotal and significant because she suggested that negotiation might serve as the main framework in learning new linguistic forms and features. Hatch (1978) further warned that the notion of negotiation and interaction may not be applicable for classroom approaches and more empirical evidences should aid in classroom decision making for the purpose of learning (Hatch, 1978). Nonetheless, a considerable amount of studies (Pica et al., 1987; Gass & Varonis, 1994; Loschky, 1994) suggest that negotiated interaction is beneficial in L2 vocabulary comprehension

and acquisition. However, the advantageous role of negotiated interaction may not be applicable to syntax acquisition as it was considered to be more difficult and complex (Pica, 1994).

Nevertheless, the role of interaction in SLA received critiques and differing opinions especially from the innatist perspective (Locke, 1690; Chomsky, 1965; Pica, 1994). Chomsky asserts humans are born with the innate ability to acquire language, known as the LAD (language acquisition device) (Chomsky, 1965). Once the LAD is activated, the child would have the mental capability to acquire and produce language. The notion of LAD disregarded the role of overt teaching and socio-cultural environment (Breen & Littlejohn, 2000). On the other hand, Krashen (1980; 1985) claims that, under any circumstantial situation, acquisition may only transpire when input is made comprehensible to the learner.

The past thirty years have seen increasingly rapid advances on the role of interaction in the field of SLA, especially on the works of negotiation (Long, 1981; 1985; Swain, 1985, Gass & Madden, 1985; Loschky, 1994; Ellis & He, 1999; de la Fuente, 2002; Shintani, 2011). Long (1981; 1985) further developed and enhanced Hatch's idea. Long (1981) investigated the discourse between non-native speakers (NNSs) and native speakers (NS) and ascertained that the NNSs and the NS had to repair their communication barrier. Long (1981) coined the term interactional modification as a direct referral to the speech repair transpired between NNSs and NS (Long, 1980; 1981; 1983). Interactional modification is known as negotiation in later years (Gass & Varonis, 1985; Larsen-Freeman & Long, 1991, Scarcella & Higa, 1981; Varonis & Gass, 1985). The term negotiation is defined as an interactional modification to achieve input or message comprehensibility between the instructor and learners, which ultimately aids the L2 learning process (Krashen, 1980; 1985). In any scenarios of miscommunication, the learners and the interlocutors would repair their

communication via various interactional modifications such as repetition, alteration and syntactical structure, substitution of vocabulary, form modification and provision of meaning to reach a state of comprehension (Pica, 1994). As suggested by Long (1980), other terms that were coined such as clarification requests, confirmation checks and comprehension check are known as repair moves.

2.2.2 L2 Vocabulary Comprehension and Vocabulary Acquisition

Several attempts have been made to establish the direct and indirect association between vocabulary comprehension and vocabulary acquisition in SLA (Long, 1983; Loschky, 1989). Loschky (1989) conducted a study to show a direct relationship between negotiated interaction, L2 comprehension and acquisition in which did not garner evidentiary support. 41 learners of Japanese as Foreign Language were subjected to an experimental study in association to vocabulary acquisition. His study revealed that there was no significant difference in relation to vocabulary acquisition and retention, although learners from the interactionally modified input attained higher level of vocabulary comprehension than learners who were given pre-modified input.

However, this attempt to establish the link between vocabulary comprehension and vocabulary acquisition was not supported by numerous researchers (Sharwood Smith, 1986; White, 1987; Ellis, 1991) on the basis that comprehension and acquisition are different language processes. The distinction between vocabulary comprehension and vocabulary acquisition will be made clear based on Sharwood Smith's (1986) definition. Vocabulary comprehension and vocabulary acquisition are known as two different methods in processing input. While vocabulary comprehension involves the need for the learners to decode a particular message that has been linguistically encoded, vocabulary acquisition encompasses the capacity of the learner to create new mental structures of new linguistics forms and meanings. During a listening comprehension task, learners have to make inferences on the linguistic indicators rather

than associating sounds to the semantics of words (Rost, 1990). In short, vocabulary comprehension is defined as a top-down process whereby learners have to internalize and address the forms in the input before the meaning. On the other hand, L2 acquisition may transpire when learners notice and pay selective attention to new linguistic forms and meanings present (Gass & Varonis, 1994; Schmidt & Frota, 1986; Schmidt, 1990). Schmidt and Frota (1986) asserted that when learners notice the gap between the input and their current linguistic knowledge, this process would facilitate L2 acquisition.

For the purpose of this study, the term vocabulary acquisition will also be defined closely based on Nation's (2001) word knowledge, which includes knowledge of form and knowledge of word meaning. Knowledge of form encompasses knowing what word looks like and what it sounds like in its phonological form. Knowledge of word meaning is divided into three parts. However, the term vocabulary acquisition will be closely linked to the first sub-division, which is form and meaning. It encompasses the ability to link the form to a meaning by forming a link between the English Language to its translation the learners' native language (Malay) (Milton, 2009).

Nonetheless, the theoretical claim that attempts to establish the link between comprehension and acquisition should not be undervalued. Ellis, Tanaka & Yamazaki (1995) mention that if learners were not able to comprehend the input provided in the first place, they would not be able to acquire new linguistic forms as their attention waned due to comprehension barrier. Therefore in 1991, Ellis proposed two views in line with the interaction hypothesis that:

- i. Comprehension does not necessarily lead to acquisition.
- ii. Only some interactional modifications that causes learner to notice promote the kind of comprehension that fosters acquisition.

Aside from that, Swain (1985) highlights that comprehensible input in SLA is not responsible in stimulating the production of output. She maintains that learners should be provided opportunities to produce pushed output in the target language to ensure that learners are elevated to a higher stage in acquiring grammatical competence. When learners are placed in a position to produce in the target language, they will be able to experiment different ways of expressing in L2 and at the same time, participate in language processing. The coined term ‘comprehensible output’ by Swain (1985) attempts to complement Krashen’s (1985) comprehensible input.

2.2.3 Input and Language Acquisition

Input in language acquisition is an essential ingredient in second language acquisition (Corder, 1967; Wesche, 1984; Krashen, 1985; Ellis, 1990; Pica, 1994). Input is referred to as the language spoken to an L2 language learner by a native speaker or a non-native speaker (Ellis, 1990). Corder (1967) iterates that input is not merely what the L2 learner listens to or offered by the interlocutor but rather the language embedded with information that goes in, making what is listen or read as as ‘input’. Krashen’s (1985) Input Hypothesis suggests that L2 learners are required to receive comprehensible input that are slightly beyond their current level of understanding in second language acquisition. However, the Input Hypothesis does not explain or describe comprehensible input in detail or how it may occur (Wesche, 1984). In response to the Input Hypothesis, numerous researches (Pica et. al., 1987; Pica, 1994) were conducted to investigate how input is made comprehensible to learners. Pica (1994) states that input is not adequate for second language acquisition to occur. Input needs to be understandable so that learners are able to reflect on how rules, forms and features of the language work. This led to investigations into pre-modified input.

Pre-modified input is described as the modification of the input designed specifically to suit the targeted learners’ current level of comprehension. Pre-modified

input is usually in forms of repetition, paraphrasing, form adjustments, syntactical modification and substitution of words (Pica et. al., 1987; Pica, 1994). Pica (1994) suggests that when input has been pre-modified to suit the learner, negotiation of meaning is not needed to facilitate L2 comprehension.

However, the idea of pre-modified input alone does not sit well with the interaction hypothesis as the latter hypothesis argues that pre-modified input limits language learners to just identifying and learning by matching the input to their current linguistic knowledge, which is not sufficient (Long, 1981). Gass and Varonis (1994) classify pre-modified input as rigid and pre-planned. To illustrate, the dialogue below depicts a form of pre-modified input:

“Teacher: What is a cow? A cow is an animal. It has four legs and eats grass. This is a picture of a cow.”

Although the input has been pre-modified to allow young L2 learners to immediately grasp the meaning of the cow, the learners are not provided the room to make negotiate for meaning produce output in the target language and commit linguistic errors. The interaction hypothesis further adds that making little or no linguistic errors inhibits the language development process (Long, 1981; 1983). Input is only beneficial for second language acquisition when learners negotiate for meaning and modify interactional structures. A rigid and pre-planned dialogue that does promote misinterpretation and negotiation is not favorable for language learning (Gass and Varonis, 1994). Therefore, learners who are engaged in interaction to negotiate for meaning are able to expedite their process of L2 comprehension. This led explorations into areas such as interaction, negotiation of meaning, the role of output and feedback in second language acquisition.

2.2.4 Interactional Input and Negotiation of Meaning

Long (1996) asserted that interaction and input are essential ingredients in second language acquisition. Input alone is insufficient for second language acquisition

and he (Long, 1996) suggested that modified interaction and negotiation for meaning between interlocutors involved in a conversation or a learning process was required for input to be made comprehensible to L2 learners. Modified interaction is defined as the modification of speech in conversation made by the native speaker or non-native speakers. This process occurs in the Interaction Hypothesis and functions as comprehensible input to language learners (Long, 1996). The Interaction Hypothesis states that comprehensible input and feedbacks on error-laden speeches helps to facilitate second language acquisition. Therefore, the steps in negotiation allow learners access to the linguistic rules, forms and feature of the target language (Bitchener, 2003).

Negotiation plays a key role in acquisition and there is substantial amount of studies (Pica et. al., 1987; Shintani, 2011; Losckhy, 1994; Ellis, Tanaka & Yamazaki, 1994) that suggest the importance of negotiation in input comprehension. Gass and Torres (2005) claim that negotiations are more applicable to lexical items than grammatical components that are more difficult. Pica (2004) encapsulates that negotiation provides opportunities to learners to modify their output and this making input comprehensible. Negotiation allows learners greater access to linguistic data.

Ellis, Tanaka and Yamazaki (1995) identify numerous benefits of negotiated interaction in L2 vocabulary acquisition. One of the benefits is that researches on negotiated interaction poses fewer problems for lexical items. Negotiated interaction may not be necessarily applicable for difficult and complex morphological elements (Sato, 1986), which may be comprehended without needing the learners to pay notice to the mentioned features. While making sense of the input received, learners may not be able to detect and identify grammatical structures embedded in problematic utterances. Vocabulary is easier to attend to and learners may seek further explanation of its meaning, which is apparent in established studies (Pica et. al., 1987; Shintani, 2011; Losckhy, 1994; Ellis, Tanaka & Yamazaki, 1994). In the earlier works of L2 vocabulary

development research, the work of Cohen (1990) looks into the various vocabulary learning strategies pertaining to memorizing and the work of Li and Hudak (1989). Li and Hudak (1989) quantitatively measure how successful learners are in making inference based on the lexical items read from texts. Elley (1989) conducted an examination on how listening to stories promotes vocabulary acquisition while Brown (1993) studied on the saliency of words in terms of verbal input delivered by the use of technology. Scant attention was directed towards comprehensible verbal input and negotiated interaction in L2 vocabulary acquisition (Ellis, Tanaka & Yamazaki, 1994).

L2 vocabulary acquisition comprises multiple areas to be explored such as the process of identifying a lexical item and its output production (Carter, 1987). Some of the widely cited researches of L2 vocabulary acquisition encompass the effects of pre-modified input and interaction towards vocabulary comprehension and acquisition (Pica, Doughty & Young, 1987; Ellis, Tanaka & Yamazaki, 1994; Losckhy, 1994; Ellis, 1995). These studies were mainly focusing on oral vocabulary acquisition, which was primarily generated from the indelible studies of language learning and communication (Hatch, 1978; Brown, 1993; Gass, 1997). Clearly, negotiated interaction and L2 vocabulary acquisition is salient for any language learning process with reference to lexical items (Zimmerman, 1997).

A lexical item is defined as a single unit or a single word in a sentence and contributes to a bigger element, which is the lexicon or vocabulary (Lewis, 1997; O'Grady, 1998). Foster's (1998) finding revealed that majority of the interactional modification that occurs in a classroom setting was emphasized on the semantics and not morphosyntax. Similarly, Pica, Kanagy and Falodun (1993) found that interactional modification that occurs during classroom negotiation was linked to vocabulary rather than L2 grammar structures. Studies on the advantageous role of interaction in L2 vocabulary acquisition are more apparent than for L2 syntax and phonological aspects.

2.2.5 Pushed Output in Negotiated Interaction

When language learners are provided opportunities to produce in the targeted language, they are required to reflect upon the syntactical structure before producing in the target language. This is known as comprehensible or pushed output, which promotes L2 comprehension and acquisition (Swain, 1985; 1995). Pushed output requires the learners to produce in the target language by experimenting with different and new linguistics forms during negotiated interaction. Therefore, in the attempt to describe the process of pushed output, Swain and Lapkin (1995) conducted a dictogloss activity, whereby Grade 8 French immersion participants were asked to think-aloud while writing and editing their short essays in French. The researchers claim that this output production provides an insight into the participants' introspective linguistic knowledge. There were 190 instances in which the participants notice a linguistic problem while trying to produce in the target language. Additionally, Swain and Lapkin (1998) explored the term 'collaborative dialogues' and recorded what the participants talk about during the production of L2, be it speaking or writing. The studies above (Swain & Lapkin, 1995; 1998) describe the processes of pushed output and that these thought processes promote L2 learning.

Pica et. al. (1989) conducted a study to explore the processes involved in pushed output as postulated by Swain (1986; 1995). The excerpt below illustrates the NNS in trying to express a message to the NS.

NNS: And one more weep weep this picture.

NS: Huh?

NNS: Another one like gun to shoot them weep weep.

NS: Oh ok ok yeah I don't have a second weapon though so that's another difference.

Conversation Excerpt cited in Pica et. al. (1989)

In the effort to ensure that the NS understood what he or she was trying to say, the NNS paraphrased the specific structure, which the NS did not understand. As can be seen, after the NNS modified his previous utterances, the NS managed to understand and decode the message expressed by NNS.

Aside from the need to supply comprehensible input, the opportunity to produce pushed output during negotiated interaction is pertinent to facilitate and promote L2 acquisition. Mackey and Philp (1998) expressed her concern that in both classroom context and the naturalistic environment, learners were more likely to observe output production rather than producing in L2. Pica's study (1992) found that there were no significant differences in terms of L2 comprehension between participants who were actively engaged in output production and participants who observed their peers in output production. This report (Pica's study, 1992) is also similar to Pica et. al.'s study (1987) with particular reference to negotiated interaction. Therefore, interaction, which facilitates input and output, is seen as a potent element in aiding L2 comprehension and acquisition. It is not obligatory for language learners to participate in active interaction because mere observation of the interaction would have profound effect on L2 comprehension.

In another study conducted by Ellis, Tanaka and Yamazaki (1994), the results of the language development between each learner were compared to the same tasks. In addressing the issue of observed interaction, the findings of the study (Ellis, Tanaka & Yamazaki, 1994) is similar to the findings of Pica (1987; 1992), whereby there was no significant difference in relation to the task scores between learners who merely observed the interaction and learners who actively interact. The results of the studies (Pica, 1987; 1992; Ellis, Tanaka & Yamazaki, 1994) revealed that actively participating in interaction might not be necessary to promote L2 acquisition. Nevertheless, active negotiation does not contribute to the negative effects of L2 development either.

2.2.6 Feedback in Negotiated Interaction

One of the important features of pushed output in negotiated interaction is feedback from the teachers (Swain, 1995; Long, 1996; Gass, 1997). Long (1996) asserts that negative feedback triggers language learners to notice linguistic features and forms that are incorrect or they may not know of. In accordance with Swain's (1985) output hypothesis, noticing language gaps requires a certain amount of attention from language learners and their developing state of mental language processors are lethal combination to aid L2 acquisition.

The Interaction Hypothesis (Long, 1980) states that in order for language learners to reap the benefits of foundational basis of interaction, learners need to notice the gap between the input and their current linguistic knowledge. This process will be based on the negative feedback received from their teachers or immediate environment provided that language learners would have time to internalize them. Similarly, Ellis (1991) proposed that linguistic awareness and assimilating are beneficial in the field of SLA. Comprehensible input facilitates the process of assimilation of the new linguistic forms to their current knowledge. If learners are not conscious, they may not notice the gap between their current interlanguage knowledge on forms and other forms that may exist in L2 (Gass & Varonis, 1994; Schmidt & Frota, 1986).

According to Gass (1991), in order to integrate the new linguistic forms into their interlanguage forms, language learners are required to pay attention or notice their linguistic gap. Such conscious processes occur through negotiated interaction (Mackey & Philp, 1998). One of the beneficial potentials of negotiated interaction is that learners are put in a situation where they need to grasp and comprehend the input received and struggle to deliver comprehensible message (Pica, 1994; Gass, 1997). Long (1996) adds that learners' selective attention and feedbacks from the teachers will be directed to the syntactical and grammatical forms as well as the semantics.

2.3 Relevant Theories as Framework of Study

There are various theories and hypotheses proposed to describe how a second language is acquired and learnt. Several of them include the Input Hypothesis as suggested by Krashen (1981), the Interactionist Hypothesis, a notion proposed by Long (1980), the Output Hypothesis by Swain (1985) and so forth.

2.3.1 The Input Hypothesis

Competing views on what is innate and what should or should not be taught dominated the field of SLA since 1960s. Chomsky (1965) pointed out that humans are equipped with an internal language processor, known as Language Acquisition Device (LAD) and that children are able to acquire language from their environment without the need for overt teaching with LAD. Krashen (1981; 1985) expanded this idea and introduced the Input Hypothesis Model that are driven by five hypotheses which are:

i. The Acquisition-Learning Hypothesis

This hypothesis focuses mainly on the definition and differences between acquisition and learning.

ii. The Natural Order Hypothesis

This hypothesis states that humans acquire and learn rules and structures of the language in progression and a certain order (Corder, 1967). The order is not determined by the simplicity of the language rules.

iii. The Monitor Hypothesis

This hypothesis places emphasis on conscious learning and language production, mainly in speaking and writing. The conscious knowledge allows human to monitor or edit their language by making corrections or alterations before producing any forms of output. However, in order for the Monitor model to be activated, the language learner must be aware of what is correct and incorrect and aware of the rules of the language. It

is worth noting that editing language forms takes a longer time to process in a study conducted by Hulstijn and Hulstijn (1984).

iv. *The Input Hypothesis*

This hypothesis looks into how acquisition is possible through comprehensible input.

v. *The Affective Filter Hypothesis*

This hypothesis describes how input is being received. The affective filter denotes a mental block due to negative emotions such as anxiety, shame, anger, defensive and so forth, which may prevent humans from using the comprehensible input to its optimal usage (Stevick, 1976). The comprehensible input may only be useful for language processing when the affective filter is lowered down.

As mentioned earlier, when Chomsky (1965) undervalued the role of overt teaching from the child's caretaker, Krashen (1985) made a clear distinction between acquisition and learning in the Acquisition-Learning Hypothesis. Both acquisition and learning refers to the child's capability in second language. The former transpires on a subconsciously while the latter projects a conscious state of recognizing and learning the language. However, in order for acquisition to occur, input received regardless of the environment must be understandable or made comprehensible based on the fourth hypothesis, which is the Input Hypothesis.

Alongside with the Natural Order Hypothesis, the Input Hypothesis argues that humans are able to acquire language structures that are one level higher than the current competence with $i + 1$. i represents our current competence in the language while 1 signifies competence that is a level higher. This may transpire with the help of additional language information, context, knowledge of our surrounding and current linguistic capability. The additional language information will be provided by the caretaker or a teacher and may be in forms of a picture or the real object.

In line with Hatch's (1978) idea, Krashen suggested that speaking or producing output in the target language does not evolve from learning linguistic structures explicitly but rather it will be developed naturally from acquiring the language. Furthermore, Krashen (1985) added the overt teaching of grammatical structures is not required. If input is made comprehensible, Chomsky's LAD will serve to process the language from the input received and internalize the rules innately, provided that the input passes through the affective filter (Slobin, 1973).

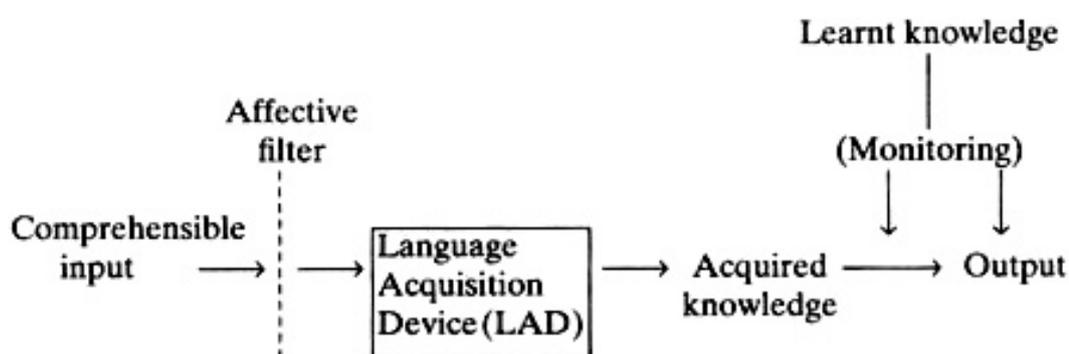


Figure 2.1: The Input Hypothesis Model of L2 Learning and Production cited from Krashen, 1982; Gregg, 1984

As illustrated in Figure 2.1, in order for people to acquire second languages, input must be made comprehensible and affective filters must be low enough to allow LAD to process the input (Krashen, 1982; Gregg, 1984). Speaking competence or producing output in the target language is a result of acquisition, which is constructed via comprehensible input. It can be concluded that comprehensible input is a crucial element for second language acquisition. Other peripheral factors may only be applicable when input is understandable and affective filters are down.

2.3.2 The Interaction Hypothesis

In 1978, Vygotsky propagated that social interaction promotes cognitive development and that language develops from the interaction between the child and the environment. In accordance with Krashen's (1981, 1983, 1985) comprehensible input and Hatch's notion on interaction (1978, 1983), Long (1981, 1996) complemented the

Interaction Hypothesis by proposing two major claims in examining interaction aids SLA. Firstly, in order for second language acquisition to take place, input or message must be made comprehensible. Secondly, input that is made comprehensible is often modified in terms of the conversational structures, which are also known as interactional modifications (Long, 1980) or negotiated interaction (Garfinkel, 1967). According to Long (1980), interactional modifications are used mainly to repair communication impairment as illustrated in Figure 2.2. While NNSs are seeking for comprehensible input to obtain meaning during a NNS and NS conversation, interactional modification codes serve as additional linguistic information to achieve message comprehensibility (Ellis, 1995). In 1978, Hatch conducted an observation by looking at the discourse between child and adult L2 learner in a natural environment. Her study revealed that the learners acquired L2 grammatical structures as a result of the interactions that they were involved in.

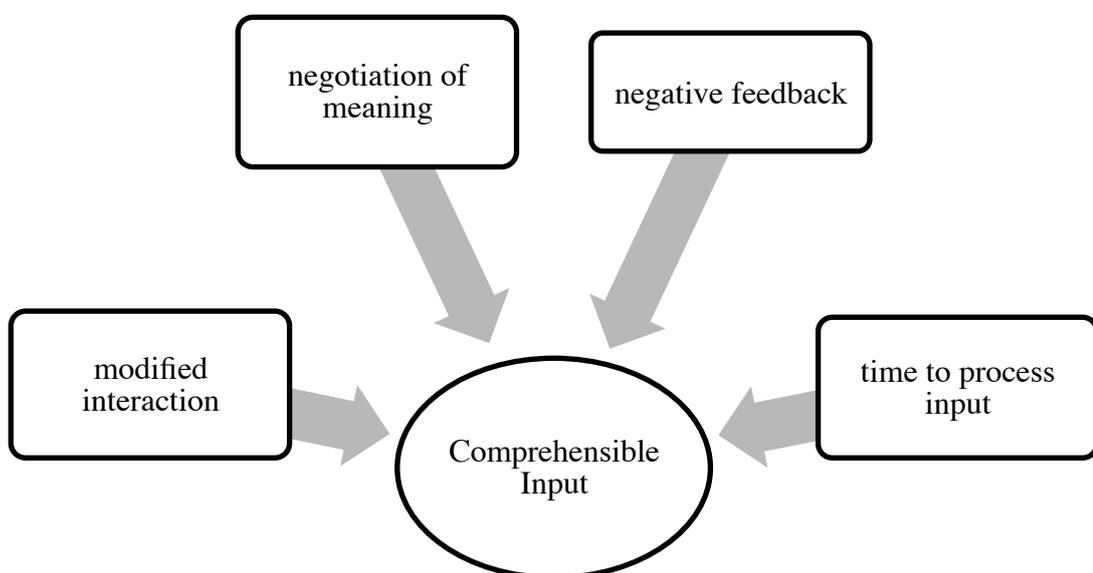


Figure 2.2: Processes involved for comprehensible input cited in Long, 1980

As discussed above, Long (1980) and Pica and Doughty (1985) categorize these interactional modifications in negotiation for meaning into four features as follows:

i. *Clarification requests*

Clarification requests refer to any forms of communication that seeks for clarification after an expression. For an example:

A: Can you pass me the turmeric powder, please?

B: What do you mean by turmeric powder?

ii. *Confirmation checks*

Confirmation checks refer to any forms of communication that seeks to confirm that the listener understood or heard correctly after an expression. For an example:

A: Mary's uncle passed away because of cancer.

B: Mary's uncle passed away? Because of cancer?

iii. *Comprehension checks*

Comprehension checks refer to the form of communication that aims to check whether the listener understood what the speaker intended to say. For an example:

A: Mary left the house. Do you get what I am saying?

iv. *Self repetitions*

There are three forms of self-repetitions; repairing, preventive and reacting. In self-repetition for repairing, the interlocutor repeats in other forms of phrases to aid communication barrier. For an example:

A: I think there should be...

B: two rats?

C: Yes, there are two rats in the house.

In self-repetition for preventive, the interlocutor repeats in other forms of phrases to prevent communication barrier. For an example:

A: Felix feels that we should share our sweets. Do you agree with Felix?

In self-repetition for reactive, the interlocutor repeats in other forms of phrases to allow the topic of discussion to flow and develop. For an example:

A: She is very clever.

B: Are you sure about that?

A: Yes, she scored above 140 for her IQ test.

Based on the examples illustrated for each feature of interactional modification, Long concludes that conversations involve higher number of conversational modification rather than alteration of input. A similar observation can be seen in based on Long's (1980) earlier work that links closely to the Interaction Hypothesis. He conducted a study on sixteen non-native speakers (Japanese) and one native speaker and witnessed reduction of syntactical complexity and variety of vocabulary used for substitution during interactions. In 1983, Long commented on the importance of comprehensible input. He suggested that input is made comprehensible through the process of negotiated interaction. Comprehensible input in bigger quantity is a necessary condition for acquisition. On the other hand, the lack of comprehensible input would result in little or no acquisition at all.

Although numerous studies (Pica & Doughty, 1985; Pica, 1987; Ellis, 1994; Ellis, Tanaka & Yamazaki, 1994) suggested that negotiated interaction promotes L2 comprehension, contradictory findings emerged for L2 acquisition (Loschky, 1994; Ellis & He, 1995). Long's (1996) Interaction Hypothesis states that the intake of comprehensible input during negotiated interaction was not sufficient to promote acquisition of new languages. Other elements such as awareness of the alterations made by the learners and negative feedback come into play with reference to the acquisition of new L2 linguistics forms such as vocabulary, morphology and syntax. These elements are further elaborated based on Swain's (1985) Output Hypothesis.

2.3.3 The Output Hypothesis

Swain (1985; 1998) broadened the scope of the theoretical framework of comprehensible input proposed by Krashen (1985) and Long's (1980) interaction

hypothesis and provided an extension of output hypothesis in the field interaction and second language learning. Kramsch (1995) and Van Lier (1996) claims that interaction not only includes the notion of comprehensible input and feedbacks but also, language learners should be provided opportunities to produce the target language (Swain, 2000). The aforementioned researchers (Kramsch, 1995 & Van Lier, 1996) asserts that the term 'output' may not be applicable for second language learning if it is confined to an information processing context. However, for the purpose of this study, the term 'output' will be based on Swain's definition that embraces the communicative process for cognitive development.

The output hypothesis is closely associated to the earlier work of Swain in 1985, whereby she conducted a study on French immersion programs in Canada. The participants are made up of children who are NNSs of French. The medium of instruction in their pre-education institution is exclusively in French. These students are given the opportunity to immerse themselves in the target language under one critical condition; the provision of an acquisition-rich environment and comprehensible input. However, under close scrutiny for six to seven years, the selected NNSs of the immersion programs showed that they were able to attain native-like reception skills but unable to attain native-like competency in speaking and writing. Based on observation, Grade 3 and 6 students were not able to engage in naturalist interaction beyond classroom-based tasks. This is primarily due to the fact that the teachers did not 'push' or provide the students the opportunity to produce in the target language. Apart from that, the teachers seldom respond to linguistic errors made by the students when they had the opportunity to produce in the target language. When teachers did address the mistakes, it was dealt unsystematically. Grammatical forms and structures were also taught covertly.

Empirical studies (Swain, 1985; 1995; 1998) seem to reach a certain level of agreement that the production of output plays a pertinent role in engaging L2 learners in deeper mental processes in comparison to processing input. When L2 learners are in control of what to produce in the target language, they are also able to apply and utilize their linguistic knowledge for communicative purposes. In order to speak or write, L2 learners have to connect grammatical and syntactical forms to semantics. While engaging in this process, they are put in position to identify inaccuracies, which is stipulated based on Swain's (1995) three functions of output illustrated in the diagram below.

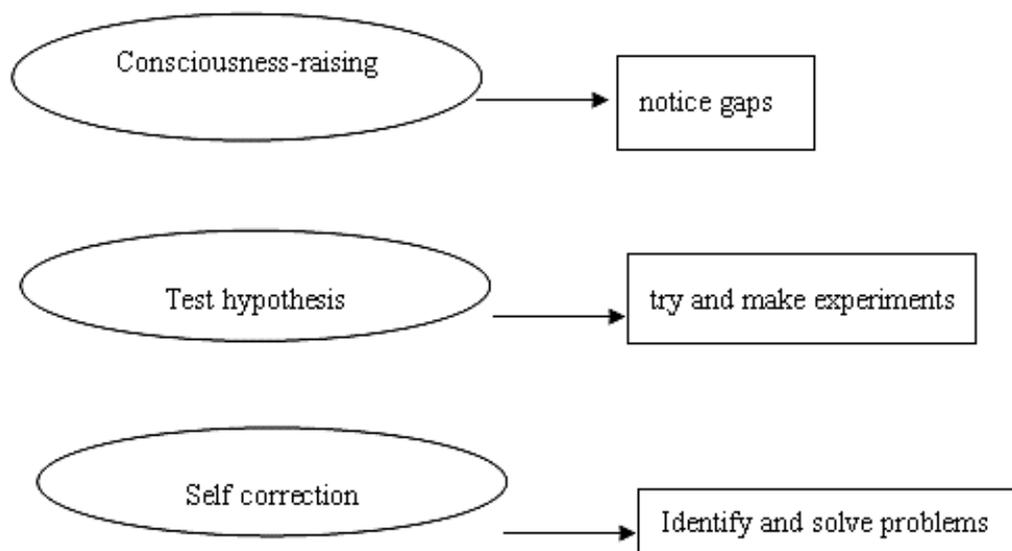


Figure 2.3: The Three Functions of Output cited in Swain, 1995

The three functions of output postulated by Swain (1995), as illustrated in Figure 2.3, encompass noticing the gap, hypothesis testing and self-correction. These functions claimed to promote language development as proposed by the output hypothesis. The first function of noticing the gap refers to how learners are able to notice the gap between their current level of linguistic knowledge and a level above their linguistic knowledge while producing the language. In this consciousness-raising process, the learners will come to a certain point when they may not be able to produce certain syntax, grammatical forms or semantics that they wish to express. Via this

productive activity, learners will realize and notice which linguistic forms they are familiar with.

The second function of hypothesis testing occurs after the process of noticing the gap. After the learners identify the linguistic units or forms, they need to scrutinize and reflect on their current linguistic knowledge. The term hypothesis testing comes into picture when the L2 learners formulate the variety of ways to express what they intend to say. By producing different structures to communicate, these learners are able to bridge their linguistic knowledge gap. This indicates that noticing the gap can trigger hypothesis testing. Swain (1995) points out that this process assists learners to attain new linguistic forms and reinforces the control over the existing language rules and knowledge.

The third function of self-correction refers to a problem-solving process that transpires between the course of noticing the gap and hypothesis testing function. It is a conscious process that allows learners to identify and reflect on how the target language works. Swain (1995) asserts that the formulation variety of ways to produce new structures in L2 is not generated based on assumptions but thorough reflection. The self-correction function facilitates acquisition through the function of noticing the gap and hypothesis testing (Swain, 1995).

2.4 Relevant Studies

2.4.1 Negotiation and Comprehensible Input

Existing literature on negotiated interaction are largely concerned with comprehensible input and L2 vocabulary acquisition (Pica et. al., 1987; Oliver 1998; Bitchener, 2003; Gass & Torres, 2005; Ng & Sheila Marina Sappathy, 2010). Pica, Young and Doughty (1987) conducted a study to examine the effects of two different input conditions in promoting L2 comprehension. Her participants consisted of twelve low-intermediate language learners. These learners were subjected to two input

conditions via a listening comprehension task that comprises 30 different directions for each picture. Similar to the studies and modifications done by Chaudron (1983) and Long (1985), the directions under the first input condition in Pica, Young and Doughty's (1983) study, labeled as pre-modified input, had alteration in terms of repetition, paraphrasing at word, phrase and sentence level and presence of conjunctions and enumerators. Three NSs, who are speakers of standard American English, were selected to deliver the directions orally to the language learners. The second input condition was not modified. However, the participants were allowed to negotiate for meaning and encouraged to ask as many questions as they would like to in order to complete the task. Below are some of the excerpts of the conversation with forms of negotiations in italics.

NS: [Initial, unmodified NS directions] Moving to the top right corner, place the two mushrooms with the three yellow dots in that grass patch down toward the road. [Directions during negotiation] Should I repeat it?

NNS: *mmm*

NS: OK. Moving to the top right corner, place the two mushrooms with the three yellow dots in that grass patch down toward the road.

NNS: *Um what means grass patch?*

NS: A piece of grass.

Conversation excerpts cited in Pica, Young and Doughty, 1987

The findings of Pica, Young and Doughty's (1987) study revealed that both pre-modified input and negotiated input facilitate L2 comprehension measured based on the number of directions being carried out successfully. The mean score for the group with pre-modified input was 69%, while the mean score for the group with negotiated input was 88% with a *t* value of 3.78, $p < 0.5$. The results indicate that, the second group, which was subjected to the negotiated input, performed significantly better than the

group with pre-modified input alone. The finding of Pica, Young and Doughty's (1987) study is supported by a similar study conducted by Pica in 1991. A similar task-based approach was employed and her study revealed that the participants, who were placed in the group with negotiated interaction, were able to comprehend the instructions well by merely witnessing their peers engaged in the process of negotiation. The results showed in Pica's (1991) study further suggest that observed negotiation promotes L2 comprehension. The conversation excerpts above show how one of the interactional modifications, which is the segmentation of language structures, aids in L2 comprehension.

While Pica, Young and Doughty's (1987) study largely deals with L2 comprehension; Ellis, Tanaka and Yamazaki (1994) conducted two experimental studies in the attempt to provide evidences that negotiated input promotes L2 acquisition, namely the Saitama and Tokyo study. The linguistic area of investigation involves vocabulary. The participants comprise 79 high school students that were in their third year in the Saitama study, whereas in the Tokyo study, the participants consist primarily of 127 high school students that were in their first year. The participants were subjected to a pretest and three posttests. In terms of the treatment, a listening comprehension task requires the participants to locate the kitchen utensils and place in the correct position of the kitchen. The differences were the type of input conditions that were given to the three different groups. The participants were randomly assigned to three groups. The control group, also known as the baseline group (B), received baseline interaction while completing the task. The pre-modified input group (PM – the experimental group) received modified instructions. This includes higher number of repetitions and the instructions were delivered at a slower rate. The other experimental group, the interactionally modified group (IM), received the baseline instructions and were allowed to interact with their teachers to negotiate for meaning. It

is worth noting that the time allocation of task for each group was not controlled. No interactions were allowed in the B group and the PM group. Based on the post-hoc Scheffé test, the IM group performed significantly better than the PM and B group in the Saitama study. There was no significant difference between the PM and B group. A similar result was reflected in the Tokyo study, the IM group outperformed the PM and B group while there was a significant difference between PM and B group. The results indicated that negotiated input and pre-modified input aid L2 comprehension. The results in the Saitama study and Tokyo study showed that the IM group performed significantly better in terms of immediate retention of the L2 vocabulary. In the Saitama study, the IM group performed significantly better than other groups after two weeks of the treatment. However, in the Tokyo study, in terms of delayed and long-term retrieval, the PM group outscored the IM group. The significance of these studies indicated that interactionally modified input might only be beneficial for short-term vocabulary retention but not long-term vocabulary retention. Accessibility to pre-modified input promotes acquisition as well.

Where negotiation for meaning and L2 comprehension and acquisition are concerned, numerous studies (Ellis, Tanaka & Yamazaki, 1994; Mackey et al., 2010; de la Fuente, 2002) are predominantly focused on secondary and undergraduate level language learners. In 1997, Ellis and Heimbach conducted a study to examine the effects of negotiated interaction in vocabulary comprehension among young learners between the ages of 5 - 6. The study revealed that only few children were voluntarily involved in negotiating for meaning, which resulted in low levels of comprehension test scores. However, Shintani (2011) argues that the cause of reluctance was principally due to the task being conducted only once. In line with Pinter's (2005) view, task repetition was necessary in Shintani's (2011) study, whereby the task was repeated nine

times. Her study revealed that task repetition led to an increase in participation in negotiation among young learners, as the task grew familiar.

Empirical evidence has consistently shown that young ESL learners are able to negotiate for meaning (Oliver, 1998; Ng & Sheila, 2010). Oliver (1998) conducted a study on negotiated interaction among 192 primary school language learners who are between the ages of 8-13. The participants were paired into 96 dyads based on their similarities in age and gender and a similar task-based approach was employed (Pica et al., 1987; de la Fuente, 2002). 48 pairs were to complete a one-way task, whereby one of the pairs would have to listen to the instruction and draw based on the instructions elicited by their respective pairs. The other 48 pairs were engaged in a two-way task, each participant in each pair would have half of the information. The pairs have to instruct each other to place the items accordingly to the positions of the kitchen based on the instructions elicited. It is also worth noting that the task was selected from commercially produced materials. The participants were audio and video recorded for two sessions and the transcripts were analyzed. Based on the mean scores tabulated, Oliver's (1998) findings revealed that children utilize self-repetition, which recorded the highest mean score of 23.98% and other-repetitions, with a mean of 23.62%. This is followed by conversational adjustments, which drops to a mean of 11.99%. These children also used less clarification requests, with a mean of 5.71%, confirmation checks, 5.72% and comprehension checks, with the lowest mean score of 0.86%. The results indicate that children were able to negotiate for meaning by employing different interactional modifications. The difference between adult and children learners, observed by Oliver (1998) is that adult learners utilizes highly on comprehension checks. To illustrate a scenario, Participant A seeks for clarification on the lexical item 'plant' by asking Participant B to spell and repeat. Participant B then provides additional information and spells the word 'plant' for Participant A to complete the task

successfully. The findings of the Oliver's (1998) study indicate that children are also able to negotiate for meaning which aids L2 comprehension.

A similar study by Ng and Sheila Marie Sappathy (2010) was investigated the effects of pre-modified input and negotiated interaction in L2 vocabulary acquisition among primary ESL learners. The participants consisted of 48 Year Five pupils from a primary national school in Malaysia. A pre-test was administered on the first day, which serves as the baseline data of the study. All the 48 participants were exposed to 14 lexical items that are concrete nouns embedded in a comprehension passage on the second day. The comprehension passage indicates an input only condition. On the third day, the intervention was given to the experimental group. 24 participants were engaged in a similar two-way communicative task as proposed by Oliver (1998). The participants were paired and were to engage in this information gap task. Each participant in a pair has a picture consisting half of the items in the kitchen. The participants would have to ask their respective pairs to spot the difference of the items available in their picture and vice versa to complete the task. This particular activity was video-recorded. On the third day, the 48 participants sat for the first post-test. The participants sat for the second post-test a week after the treatment. The final post-test was administered three months later. The findings indicate a significant difference in the vocabulary retained for both groups. However, the study revealed that the group with pre-modified input and negotiated interaction performed significantly better, with a Partial Eta Squared value of 0.925, than the group with pre-modified input only, with a value of 0.667. It is also worth noting that the group with pre-modified input and interaction yielded positive results for immediate and delayed vocabulary retention. Additionally, the mean value for Post-test 1 (7.271), 2 (8.042) and 3 (7.417) were significantly higher than the pre-test (2.521). The mean difference value for Post-test 2 (one week later – delayed vocabulary retention) at 5.521 is significantly higher than

Post-test 1 (a day after), while Post-test 3 (three months later) showed no significant difference in comparison to Post-test 2. The results indicated that learners who took part in the information gap task attained higher levels of vocabulary acquisition and retention. It is worth noting that young ESL learners were able to negotiate to achieve message comprehensibility.

In 2005, Gass and Torres expanded the ideas of previous researches (Pica et. al., 1987; Oliver, 1998; Ng & Sheila Marie Sappathy, 2010) by including different linguistic elements such as syntax. The researchers conducted an experimental study on the effects of separating and combining interaction with input on different L2 areas. The participants comprised 102 L2 Spanish learners who were in their third semester. The participants were subjected under four different conditions: (a) input-only materials, (b) interaction-only materials, (c) input-only materials preceded by interaction and (d) interaction-only materials preceded by input. The two tested language areas were grammatical structures, which primarily consist of gender agreement and the *estar* copula, and vocabulary items. In the input-only material condition, the participants were exposed to the lexical items with their corresponding adjectives and locations from listening activities. The objective of the input-only activity is to identify the personal belongings that were stolen from the hotel room. In the interaction-only material group, the participants were engaged in two communicative tasks; a jigsaw and information gap task. Similar to Oliver's (1998) task, the participants were paired up and were given the same picture with half of the information for each pair. Participant A and B in a pair would have to spot the differences in their respective pictures by negotiating of meaning. A part from identifying the items, colours and location of the items were also part of the tested L2 area. This task itself tests three items, which are nouns, adjectives and the *estar* copula. In the information-gap activity, the participants were required to instruct the researcher to draw based on the pictures provided to the participants. The

study (Gass & Torres, 2005) also emulated a pre-test and post-test design as described in Ng and Sheila's study (2010). The analysis of the data revealed that in terms of all the experimental groups, the group that showed significant difference in relation to vocabulary is the combination of interaction followed by input group. According to Gass and Torres (2005), the positive results were due to the lexical items used in the study, which consisted of concrete nouns. By incorporating the use of the participants' linguistic knowledge, interaction is beneficial for vocabulary learning. Conversely, the researchers mentioned that input-only materials might be more beneficial for more complex linguistic forms such as syntax as more considerations and attention should be paid to internalize these structures.

2.4.2 Negotiation and Comprehensible Output

In 2002, de la Fuente contributed to the growing body of literature in negotiated interaction and comprehensible output. She examined the role of different type of interactions on L2 vocabulary acquisition. de la Fuente (2002) modeled her study after Ellis and He's (1999) study, which looks at three dependent variables, which are L2 vocabulary comprehension, receptive vocabulary acquisition and productive vocabulary acquisition. In this experimental study, 32 Spanish L2 university students were randomly assigned to three different treatments; pre-modified input (NNPI); negotiation of input without output (NIWO); and negotiation of input plus output (NIPO). This research adopted the pre-test and post-test design. On the first day, receptive and productive Vocabulary Knowledge Scale (Wesche & Paribakht, 1996) test were administered to all participants. A listening comprehension task, which serves to measure vocabulary comprehension, will be administered on the second day. The task required the participants to identify the lexical items and placed it accordingly on the numbered sheets based on the instructions. The NNPI group was not allowed to interact while the NIWO group was allowed to interact. The task is repeated. In the second

attempt, the NIPO group will assumed reverse roles and instructed the instructor to place the lexical items on the numbered sheets accordingly. Another similar task was conducted on the third day. Three post-tests were administered a day after the treatment to check on immediate vocabulary retention, a week later to check on delayed vocabulary retention and three months later to check on long term vocabulary retention. Similar to Ng & Sheila's study (2010), her (de la Fuente, 2002) study revealed that groups (NIWO + NIPO) that were involved in negotiated interaction, with a *t*-value of 3.967, outperformed the NNPI group in relation to vocabulary comprehension. In terms of receptive acquisition, there was no significant difference between the NIWO and NIPO group (3.28 vs. 4.97, $p = .129$). Nevertheless, the NIWO (1.11 vs. 3.286, $p = .054$) and NIPO (1.11 vs. 4.97, $p = .001$) group attained higher level of receptive acquisition than the NNPI group. Additionally, the NIPO group outperformed NIWO and NNPI in terms of productive acquisition and acquiring more words in delayed vocabulary retention. The results from de la Fuente's study and the current study may provide support to Swain's output hypothesis (1985) where comprehensible input is not sufficient for L2 acquisition. In order to facilitate L2 receptive and productive vocabulary acquisition, learners should be given the room to interact and produce in L2.

Mackey (2010) conducted a detailed examination of input, negotiation and pushed output. An important departure of her study from others (Pica et. al., 1987; Ellis, Tanaka & Yamazaki, 1994; Ellis & he, 1995; de la Fuente, 2002; Shintani, 2011) is that the language area investigated revolves around grammatical question forms and structures. The participants were 34 adult ESL learners from a private English school in Australia. These participants were then randomly assigned to five input conditions, with one control group and four experimental groups. The first experimental group was labeled as the Interactors. They were encouraged to interactionally modify the input for meaning. The second experimental group was labeled as the Interactors Unreadies. This

group received similar input and was encouraged to interact. The only difference is that their developmental level is lower than the first group. The third experimental group was known as the Observers. They were not allowed to actively interact but could only observe the interactionally modified input. The fourth experimental group was labeled as Scripteds. They were exposed to pre-modified input. The control group received no treatment prior to the post-tests. The tests would be in forms of 'Spot the Difference test' while the treatment consisted of three activities; picture-drawing task, story completion task and story sequencing task. The study revealed three significant findings, which were relevant to the current study. The results, however, were contradictory to Pica et. al.'s study (1995) where participants who actively participate in the negotiation process benefited second language development in comparison to the participants who observes. Without undervaluing the role of observed interaction, 98% of the participants in the Observers group managed to complete the task based on the input receive. This indicates that observed interaction has some limited effects to L2 development. The input received came in two forms; the instructions elicited by the NS instructors and the negotiation between the NS and the NNSs. Finally, in terms of time effects, the Interactors group showed significant increment in terms of production of questions in the posttests that were administered a week after and a month after the treatment. This may be due to the fact that more advanced structures need to be acquired over time through the process of internalization. Thus, practicing through repetition and production is proven to be more beneficial for complex grammatical structures in L2.

In a recent study by Shintani (2011), a quasi-experimental study investigated the effectiveness of input-based task and productive-based task on L2 vocabulary acquisition. Similar to de la Fuente (2002) and Ellis, Tanaka and Tamazaki (1994), the vocabulary items were selected as these studies mentioned support that vocabulary is a

better choice when dealing with participants who possess limited knowledge in the L2. Another relevance of this particular study (Shintani, 2011) to the present study is that the participants consisted of 36 young EFL learners who are the age of 6-8. They were randomly assigned to three groups. The first experimental group, which is the input-based group (IB), was not required to produce in L2. The second experimental group, which is the productive-based group (PB), was required to produce in L2. The third group, which served as the control group, was exposed to any of the tested items. However, both experimental groups were given the opportunity to receive input and produce output. The IB group performed three listen and do tasks, which are 'Help the Zoo and the Supermarket', 'Help the Animals' and 'Listening Bingo Game'. The tasks were repeated. The PB group performed five tasks, which are 'Listen and Repeat', 'Guess the Hidden Items', 'Throwing Dice', 'Production Bingo Game' and 'Kim's Game'. After the treatment, all the participants were subjected to four assessments whereby two of the tests measured comprehension and the other two tests measured production. A pre-test was administered two weeks before the treatment and two post-tests were carried out two weeks and four weeks after the treatment. The results of the study revealed that input-based tasks enabled young L2 language learners to acquire L2 vocabulary. There were significant differences between the IB group and the control group in relation to receptive and productive acquisition based on the means of Posttest 1 and Posttest 2. Additionally, the productive-based tasks chosen proved to be beneficial for L2 vocabulary acquisition among young L2 learners. The PB group outperformed the control group in relation to Posttest 1 and Posttest 2. While comparing the performance of the IB and the PB group, there were no significance differences in terms of the scores of the three tests. However, in one of the tests (Category Task Test), the IB group outscored the PB group. Although the PB group had to produce in L2 based on the teacher's instructions, the IB group experienced higher level of interaction

in a naturalistic manner. Similar to Ng and Sheila's (2010) study, young ESL learners are able to produce positive results in L2 vocabulary acquisition when they are given the opportunity to produce in the target language.

It has also been suggested that there was no significant difference in terms of the different oral interactions between peers and peer-teacher instruction in L2 acquisition (Van de Branden, 1997). In 1997, Van de Branden explicitly examined the effects of different types of oral interactions on learner's productive knowledge among 48 young language learners. Slightly more than half of the participants included NSs of Dutch and the rest, NNSs of Dutch. The research design adheres to a typical quasi-experimental study where the participants were randomly assigned to three groups of 16 under three different conditions; (a) peer-to-peer interaction, (b) peer-teacher instruction and (c) no pushed output, which directly serves as the comparison group. For the treatment, the participants are required to solve a murder case by either drawing or instructing their partner to draw 2 out of 4 suspects of the murderers. For the peer-to-peer interaction condition, 8 dyads were formed and their partners were their peers. One will assume the role of the listener and the other, the describer. On the other hand for the peer-teacher instruction, each participant will be paired with the teacher. The participants for this group assumed the role of the describer where they have to describe the physical appearances of the 2 suspects and the teacher would have to draw based on the instructions. For the third condition, the participants in this group were not required to produce any output and no interaction will take place. After the treatment, the participants solved the murder case. The data were quantitatively and qualitatively analyzed. However, for the purpose of this study, only the quantitative portion will be reviewed. Based on the results of the pre-test and post-tests analyses, the study revealed that children interactionally modify and produce output regardless if the feedback received was from their peers or teachers. Instead, the type of feedbacks received

determined the learners' output. The feedbacks may have affected their output production for long-term vocabulary retention, as the results indicate no significant differences between the immediate post-tests' scores. Additionally, learners who are engaged in interaction with partners produced greater amount of output significantly than learners who are not involved in any forms of negotiation as shown by the comparison group. However based on the post-tests' scores, negotiated interaction may have limited or no effects upon the L2 grammatical morphology.

2.4.3 Task-Based Communicative Activities for L2 Vocabulary Acquisition

Previous studies have reported the indirect relationship between the role of task-based lessons and output production in L2 vocabulary acquisition (Gass & Varonis, 1985; Bitchener, 2003; de la Fuente, 2006). In an experimental study, de la Fuente (2006) examined the role of three types of lesson in promoting L2 vocabulary acquisition, which are task-based lesson, form-focused instructions and the traditional Presentation, Practice and Production lesson (PPP). For the PPP lesson, the participants are exposed to the same input of the dialogues listened on the first day in the presentation stage. In the practice stage, the participants read the dialogues in pairs and fill in the blanks with the tested items. At the production stage, the participants had to role-play in pairs in a restaurant setting and not restricted to the existing script. For the task-based lesson, the participants are exposed to the same dialogues for the first stage. Then, the participants have to role-play in pairs as a customer and a waiter focusing on forms and meaning. It is in an information gap format. Finally, the participants have to work in pairs to prepare a perfect menu and justify the reasons why the menu was chosen. For the form-focused lesson, the procedures are similar to the task-based lesson for the first and second task. However, the third task would engage instructor's role where the instructors would address linguistic issues arise from the tasks. Post-test 1 will be administered after the treatment and Post-test 2 will be administered a week

after the treatment. The findings of the study revealed the task-based lesson group scored significantly higher than the PPP lesson group. This indicates that task-based lessons is proven beneficial for L2 vocabulary acquisition based on the Post-test 2 mean scores that suggests delayed vocabulary retention. de la Fuente (2006) argued that this is primarily due to the fact that task-based lesson provide learners with more opportunities to interact and producing in L2.

Similarly in a small-scale study, Bitchener (2003) reaches conclusion that in order to facilitate delayed and long-term vocabulary retention, the task conducted should be subjected to two conditions. Bitchener's study revealed that repetition of task was necessary and information gap format was a critical element in a decision-making or a two-way communicative task to aid the vocabulary acquisition especially where concrete nouns are concerned. Aside from that, communicative tasks ensure that young and adult language learners alike to produce minimal interaction in order for the treatment to take place. However, it is found that the repeated information gap tasks have limited effects on other linguistic aspects such as abstract nouns and adjectives. Similar to the studies attempted by de la Fuente (2006), it is recommended that teachers employ task-based lessons in any of L2 vocabulary development activities. Numerous researchers (Pica et. al., 1983; Nunan, 1989; Lee, 2000) propose that some of the communicative tasks that provide opportunities for negotiation are information gap activity, decision-making, jigsaw task and problem solving. The results from two studies, Long (1983) and Doughty and Pica (1986), were similar. Both studies revealed that there were significant differences in the different types of tasks and the number of interactions produced. Based on the results of their studies, two-way communicative tasks (two-way information gap task) yielded more numbers of modified interactions than one-way interactive tasks and decision-making tasks. However, Gass and Varonis' (1985) study did not support the works of Long (1983) and Doughty and Pica (1986).

The researchers (Gass and Varonis, 1985) examined the effects of a one-way task, a two-way task and the frequency of interactional modifications transpired. The results of the study revealed that there was no significant difference in terms of the frequency of the interactional modification produced by the NNS/NNS interactions. Nevertheless, Gass and Varonis (1985) proposed that differences in the task designed and built-in caused such comparisons to be rendered difficult.

2.4.4 Learning Vocabulary using Pictures

Numerous studies have attempted to explain how pictures are easily remembered in comparison to words as demonstrated in memory research (Paivio & Csapo, 1973; Paivio, Rogers, & Smythe, 1968). The findings from these studies revealed that the participants performed significantly better in recalling the names of the pictures than word labels. This effect is known as picture superiority effect (Paivio, 1971). Paivio (1971) conducted a study on picture superiority effect in his dual-coding theory. His work (Paivio, 1971) suggests that pictures are better in triggering the memory than words because pictures encompasses verbal and image codes. The study is supported by the findings from Snodgrass, Wasser, Finkelstein, & Goldberg (1974), whereby the participants found it easier to name the pictures rather than recalling the words without the help of pictures. Picture superiority is also in line with the works of Craik and Lockhart (1972). Craik and Lockhart (1972) introduced the levels-of-processing theory, which suggested that pictures are easily etched in memories due to the inclusion of semantics (Nelson, Reed, & McEvoy, 1977). Research evidences from a study by Potter and Faulconer (1975) substantiated the levels-of-processing theory by demonstrating that pictures are categorized at a faster rate than words. In 1976, Nelson, Reed and Walling put forth the sensory-semantic model, which argues that picture gained greater advantage because pictures appeal to the visual sensory rather than words. Nelson, Reed and Walling (1976) supported the above model by displaying that

the picture superiority effect may be eradicated by incorporating visual stimuli in the pictures used for word association. The above studies helped to establish the importance pictures as an effective vocabulary-learning tool in classrooms and substantiate the selection of pictures used in the listen and do task for primary ESL learners of this present study.

Zooming into the second language learning and foreign language-learning environment, numerous scholars (Herron, Hanley, & Cole, 1995; Mueller, 1980; Omaggio, 1979) suggest incorporating the use of pictures into language learning materials due to its positive effects. Taglieber, Johnson and Yarbrough (1988) reported that in order to comprehend the meaning of the word '*der Hund*' in German, German language learners are required to connect the word to the notion of a 'dog'. It may be done in two different ways; providing a direct English translation of the word dog or associating the word '*der Hund*' with a picture of a dog. Reports (Hudson, 1982; Taglieber, Johnson, & Yarbrough, 1988) have shown that picture association has positive effects on vocabulary learning in comparison to direct translation.

In a language-learning context, Harmer (2001) asserts that language teachers often rely on picture or images that are widely available in newspapers, magazines and books to promote vocabulary learning and acquisition. The use of pictures, it has been claimed, makes language-learning processes pleasant and unforgettable (Hatch & Brown, 1995). Pictures were said to capture young learners' attention and help to maintain motivation throughout the entire lesson by easing comprehension of the vocabulary learnt (Carpenter & Olson, 2012). Harmer (2001) adds that direct reference of concrete nouns to pictures makes it easier to recall vocabulary and to check for meaning. For an example, in order to explain to meaning of an aeroplane, a picture of an aeroplane will sum up the meaning naturally as young learners make direct association of aeroplane to the real object.

Nation (1990) proposed a variety of rudimentary methods for language learners to learn new vocabularies, specifically targeting on young ESL learners in formal classrooms. One of the fundamental techniques mentioned would be the use of pictures from widely available mass media such as pictures from the Internet, photographs, drawings on the board and pictures from books. Harmer (2001) recommended that the suitability of the pictures is dependent on the language, the age and level of the language learners. According to Carpenter and Olson (2012), animated cartoon, comic strips and cartoon still pictures have also shown to have positive effects of language learning among young language learners. However, if the pictures used for older learners are deemed as childish, language learners may be put off. The size of the picture is one of the important elements in the selection process. A big picture is necessary for a whole class presentation, a medium-sized picture is appropriate for group activity while a small picture is applicable for individual and pair work activity (Harmer, 2001). Therefore, the selection of a cartoon still pictures have been chosen to depict the nouns used in this present study to appeal to and maintain motivation among young language learners. The pictures printed are small as it is deemed suitable for individual activity.

2.4.5 Contradictory Findings

In other parts of world, literature has emerged that offers contradictory findings about negotiated interaction and comprehensible output (Ellis, 1985; Loschky, 1994; Ellis & He, 1999). In a study conducted by Loschky (1994), contradictory findings were displayed when he attempted to empirically experiment whether input and interactionally modified input promotes L2 comprehension and acquisition. The participants are made up of 41 undergraduate students who are studying Japanese as Foreign Language and are randomly assigned to three different groups, which are; (a) baseline input group (BL), (b) pre-modified input group (PM) and (c) baseline input

group (INT). It is important and interesting to note that, similar to studies done by Gass and Torres (2005) and Mackey et al. (2010), Loschky ventured into another turf of linguistic aspects, vocabulary items and Japanese locative sentence structures. The study follows a pre-test and post-test design. For the treatment, the participants have to complete three sets of listening tasks that are known as Still Lifes, Maps and Shapes. Still Lifes involves the arrangements of objects while Maps requires the participants recognize locations pictured in a map with landmarks like temples, trains, parks and so forth. Shapes essentially encompass arrangements of two-dimensional images such as squares, rectangles, triangles and so on. For the BL group, the instructors would read out instructions with simple sentence structures. The instructions were not modified in any ways and the participants are not allowed to negotiate for meaning with the instructors or their peers. The delivery rate of the instruction is at a normal rate. For the PM group, the instructors would read the same baseline input with additional sentences that contains, elaboration, repetition and clarifications of the baseline instruction. The instructors are allowed to modify their speed of delivery and stress certain words while reading the instructions. Similar to the BL group, no interaction is allowed for the PM group. In the INT group, the instructors will deliver the baseline instructions. Negotiation for meaning between the learners and the instructors are allowed which includes, repetitions, paraphrasing, clarification requests and others. Time for the tasks are not controlled for all groups. His study revealed that although learners from the interactionally modified input attain higher level of vocabulary comprehension than the BL group ($t(df == 36) == 2.281, P < .05$) and the PM group ($t(df == 36) == 3.227, p < .05$). , there were no significant difference in relation to vocabulary acquisition and retention. The difference observed of this particular study from the rest (nouns) was the choice of tested lexical items, which is the Japanese locative expression.

In 1999, Ellis and He conducted an experimental study to determine the

effectiveness of modified input, negotiated input and the need for output production on L2 vocabulary comprehension, receptive vocabulary acquisition and productive vocabulary acquisition. The participants comprises of 50 undergraduate students who are of intermediate level proficiency in English. The participants are also assigned to three groups, which are the Pre-modified Input Group, Interactionally Modified Input Group, and the Output Group. The study adhered closely to the pre-test and post-test design; (a) one pre-test administered a week before the treatment, (b) Post-test 1 administered a week after the treatment, (c) Post-test 2 administered two weeks after the treatment, (d) Post-test 3 administered three weeks after the treatment and (e) Post-test 4 and 5 administered four weeks after the treatment. For the treatment, each group will have to perform the same task under three different conditions. All participants are required to locate and place the furniture accordingly to the correct location in the picture of an apartment provided. This task will also serve to measure vocabulary comprehension and therefore, no separate measures for comprehension will be administered. In the pre-modified input group, the directions were modified with insertions of repetition and definitions of the lexical item tested. The instructions were delivered at a slower rate of 90 words per minute. In the interactionally modified input group, the baseline instructions were given to the participants. The delivery rate was 180 words per minute and interactions were allowed. For the output group, the participants are to produce 10 directions for each piece of furniture and perform the same tasks mentioned earlier with their respective pairs. The findings of the study revealed that the modified output group scored significantly better than the other two groups at $f = 6.31$, $p < .01$ in terms of vocabulary comprehension. In relation to receptive vocabulary acquisition, the output group outscored both input groups significantly. However, there was a significant difference between both input groups. This indicates that the output group and the interactionally modified input group

performed significantly better than the pre-modified input group. Finally, the output group outperformed the two input groups in terms of productive vocabulary acquisition. This study confirmed that when learners are provided the opportunity to produce output, they gain greater advantage in vocabulary comprehension, receptive and productive vocabulary acquisition. Pre-modified input and interactionally modified input may have limited effects in L2 vocabulary comprehension and acquisition. However, de la Fuente (2002) argued that this might be primarily caused by one of the uncontrolled variable, which is the time allocated for each task. A similar findings by Ellis (1995) further substantiated that negotiated input may not be as efficient as pre-modified input. For this particular study, time was a controlled variable where one word is acquired within a minute of input. The findings of the study revealed that negotiated interaction promotes L2 vocabulary acquisition when language learners are pushed to recognize unknown words within the input. Nevertheless, to identify what transpired within the negotiation that instigated the learners to notice the unknown words and acquire the vocabulary was not mentioned in the study.

2.5 Conclusion

This chapter has presented a relevant discussion of the different oral interactions; pre-modified input, negotiated input and output production in L2 vocabulary comprehension and acquisition, theoretical frameworks on negotiated input and output that the present study is grounded on and previous studies that are directly and indirectly associated to the present study. An experimental study to determine the effectiveness of each oral condition is necessary which may affect language-learning effectiveness in the classroom. The study is conducted to ensure that primary ESL learners would receive effective treatments and measures to enhance their communication skills. The acquired skills would be able to enhance pupil learning and equip them with the required skills for the future.

Various previous studies were presented on how to conduct the experiment and the criteria that should be considered while conducting the experiment. Studies have shown that an experimental measure would be most effective to ascertain the effectiveness of each condition. During the treatment, an information gap formatted task is required to ensure suitability and that interaction takes place among young ESL learners. Finally, pictures that are deemed suitable and colourful are used to capture young learners' attention and maintain their motivation.

CHAPTER 3: METHODOLOGY

3.1 Introduction

This study intends to investigate the effects of negotiated interaction and output in L2 vocabulary acquisition among primary ESL learners in Kuala Langat, Selangor. In addition, variables, which could influence the findings and predict the effectiveness of negotiation and output production among the pupils, are identified. In order to obtain valid and reliable data to explain the phenomena, systematic data collection procedures were adhered to. This chapter will present the procedures used in this experiment. It is subdivided into research design, participants, measures, operational definitions of variables, procedures, data analysis and pilot study analysis.

3.2 Research Design

This research is quantitative in nature. Quantitative method encompasses data collection in such a way that information is derived from numbers and subjected to statistical treatment to explain a phenomenon investigated (Aliaga & Gunderson, 2000; Creswell, 2005). Creswell (2005) further asserts that quantitative research may help to support or disprove knowledge claims. With that in mind, the underlying premise that frames this current study is based on the positivist paradigm. Traditionally, positivists employ a systematic and scientific approach in their studies. Hughes (2001) suggests that phenomenon can be explained through fixed and universal knowledge laws. Some of the key components in this paradigm are observation and collection of data, tracing for patterns and formation of theory, formulation of hypothesis, conducting studies to test hypothesis and support or adjust the theory.

Therefore, it was decided that an experimental design, modeled after de la Fuente's (2002) experiment, is an appropriate method to test negotiated interaction and the output hypothesis in L2 vocabulary acquisition among young ESL learners in Kuala Langat, Selangor. One of the principle advantages of experimental design is that the

procedures are replicable as it is clearly laid out as explicitly described in de la Fuente's (2002) study. The same experiment may need to be conducted numerous times to arrive at a universal principle (Federer, 1955; Kirk, 1982). Additionally, the research objectives of the experiment provide a focused scope that may be independent of the researcher's personal judgment and bias, which may influence the results of the study (Keppel & Wickens, 2004; Keppel, 1991). Finally, this experiment allows the researcher to determine the relationship between independent variables investigated towards the dependent variables. The researcher will have higher degree of control of the variables.

The treatments would be administered to two experimental groups during the experiment. Data sets from the pre-test and post-tests from the experiment would be measured and analyzed through SPSS (Statistical Package for Social Sciences). SPSS is a statistical tool that helps in the process of data and makes informed decisions (IBM, 2014). The scores from pre-test and post-tests will serve as the primary data to compare measureable changes between groups from the intervention (Huck & McLean, 1975; Maxwell, Delaney & Manheimer; 1985).

3.3 Participants

Abdul Fatah and Mohd Majid (1993) laid out four procedures in selecting an appropriate sample for a study as stipulated below:

1. Define your population
2. Obtain a complete and accurate list of the population
3. Select a representative unit from the list
4. Obtain an adequate size of sample to represent the traits of the population.

Target population is identified as a group of people whom are able to provide information and of interest to the researcher (Frankael & Wallen, 1996; Levin & Rubin, 2000).

Table 3.1: Sample Size Based on Stratified Random Assignment for a Primary National School in Kuala Langat, Selangor

Groups		Number of Samples	
		Gender	Total
GIPO (Control Group)	Female	8	15
	Male	7	
GINW (Experimental Group)	Female	8	15
	Male	7	
GINP (Experimental Group)	Female	8	15
	Male	7	
Total		45	45
Percentage		100	100

As illustrated in Table 3.1, the target population comprised of Primary 3 pupils who are currently studying in national primary schools in Kuala Langat, Selangor. The samples will be chosen based on random sampling method (Creswell, 2005; Bryman, 2012), whereby the participants are made up of representative sample of 45 Primary 3 pupils from a primary national school in Kuala Langat district in Selangor. The participants were selected on the basis of geographical location, which is the suburban area of Telok Panglima Garang. A suburban area was chosen because the participants lack language-learning opportunities out of classroom settings (Samuel & Bakar, 2008). An informed consent for the experiment was obtained from each participant's legal guardians (see Appendix L1 and Appendix L2). The researcher informs the legal guardians of the objectives and procedure of the study. Respective school authorities were informed prior to the experimental study (see Appendix M).

The samples will be randomly assigned to three groups (Creswell, 2005; Bryman, 2012) with a recommended sample size of 15 participants for each group (Gall et al., 2003; Creswell, 2005; Bryman, 2012). Each group has an equal variation of 8 female participants and 7 male participants. However, it is worth noting that in order to

achieve true experiment and guard threats against internal validity, the selected participants must be randomly assigned to a control group and two experimental groups (Johnson & Christensen, 2000) for internal validity. The term random selection or sampling and random assignment should not be confused. Random selection transpires at a macro level, whereby a representative sample of participants is selected from a larger population (Frankael & Norman, 1996; Levin & Rubin, 2000). While in random assignment, all of the selected participants would have an equal chance to be assigned to any groups (Trochim, 2006). For this study, a randomizer software, QuickCalcs (2013), is used to randomly assign all the participants. Firstly, the researcher keys in 21 male participants into the randomizer software (QuickCalcs, 2013). Then, QuickCalcs distributes the male participants into three groups, with 7 participants in each group. The steps are repeated for 24 female participants, with 8 participants in each group.

Based on Johnson and Christensen's (2000) proposed sampling technique for classroom experiments, extraneous variables of all groups should be highly controlled to guard against validity threats to the findings such as age, gender, level of proficiency and so forth. Essentially, extraneous variables are defined as noise variables or unfavorable factors that are present, which would influence the causal relationship of the variables in a particular study (Hall, 1998). All participants possess similar ethnic background and mother tongue, which is Malay. This information is verified through the English Language Background Questionnaire (see Appendix A1 and Appendix 2), which is distributed to the participants' parents or guardians. The questionnaire is prepared in two languages (English and Malay) and the participants' parents or guardians may fill in the questionnaire in the language that they are comfortable in. In terms of their academic background, all participants received formal pre-education at the age of five and primary education in primary national schools with Malay as the

medium of instruction. The participants experience formal learning in English for 210 minutes per week as set by the Ministry of Education, Malaysia (2010).

The selection of participants in terms of age is closely linked to their language proficiency. All of the participants are 9 years of age. The participants possess average proficiency level in English, which is measured based on the standard performance band provided by the School-based Assessment Management System (SPPBS) created by the Ministry of Education in Malaysia. The participants are currently placed in Band 2, whereby they are able to apply word attack skills and able to spell common sight words (see Appendix B). The acquired skills in Band 2 enables these participants to accomplish the tasks designed in the experiment as recommended by Shintani (2011). Other form of measurement includes their monthly test scores in English Language subject, which is an average of 50-55 percent. The exclusion of other ethnicities, different ages and different levels of proficiency are done so that independent variables of interest may be clearly examined for cause of the change in the experiment (Rosenthal & Rosnow, 1991).

3.4 Operational Definitions

The participants ($n: 45$) were divided into three groups respectively; one control group (GPIO) and two experimental groups (GINW and GINP) based on the independent variables of interest in this study as illustrated in Figure 3.1. The group under GPIO will serve as the baseline group and receive neutral treatment in order to determine significant changes of the intervention, which is pre-modified input only (Creswell, 2005; Bryman, 2012). On the other hand, additional interventions that are under investigation will be subjected to the groups under GINW and the GINP. The GINW group receives pre-modified input with negotiated interaction, while the GINP group receives pre-modified input with negotiated interaction and produces output.

Pre-modified input is described as the modification of the original instruction

elicited by the interlocutors. The term negotiation is defined as an interactional modification to achieve input or message comprehensibility between interlocutors and learners (Krashen, 1980; 1985). Pre-modified input and negotiated interaction usually transpires in forms of repetition, paraphrasing, form adjustments, syntactical modification and substitution of words (Pica, 1994). The term output production proposed by Swain (1985), supported by recent works of Long (1996), claims that learners should be provided with a room for producing the new vocabulary learnt to increase productive vocabulary acquisition.

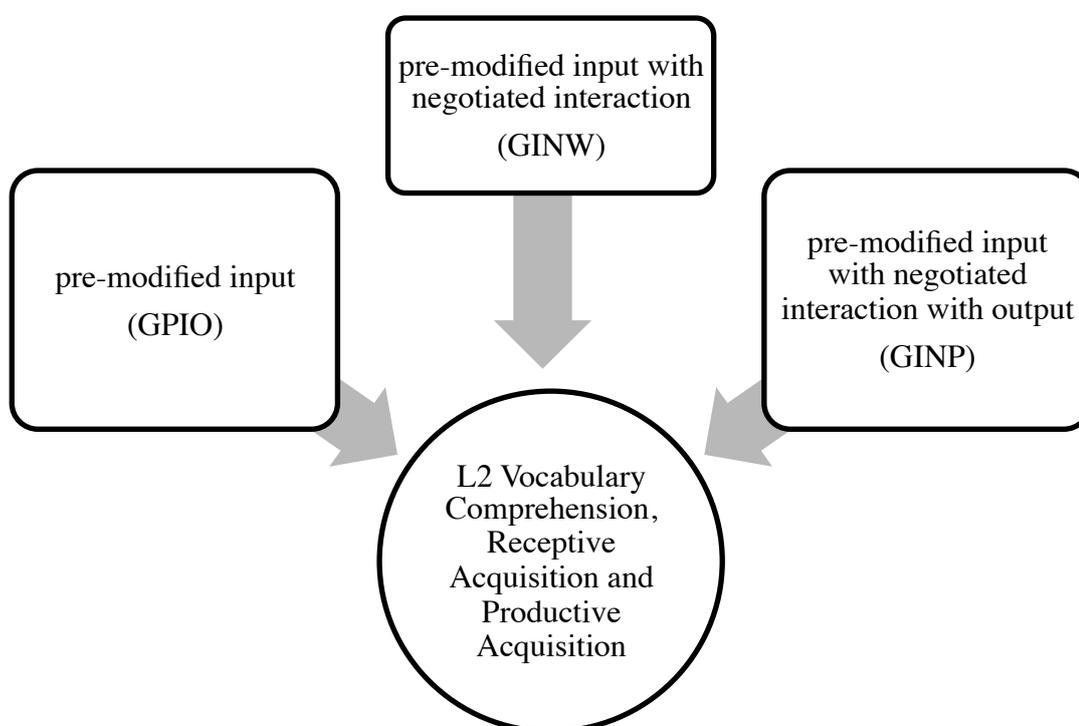


Figure 3.1: Independent and Dependent Variables

An experiment has been conducted to measure the effectiveness of the three independent variables discussed above towards three dependent variables, namely, L2 vocabulary comprehension, L2 receptive vocabulary acquisition and L2 productive vocabulary acquisition. The term L2 vocabulary comprehension is defined as the ability to understand the meaning and recognize a particular lexical item orally (Swain, 1985). Apart from that, the distinction between receptive and productive vocabulary acquisition will be based on Melka Teichroew's (1982) definition. He suggested that

receptive and productive acquisition constitutes a continuum. Receptive acquisition steadily progresses towards productive mastery, due to the complexity of productive tasks (Faerch et al., 1984; Palmberg 1987). Receptive and productive vocabulary acquisitions are primarily advantageous for lexical learning and less beneficial for grammatical morphology (de la Fuente, 2002). In a language-learning context, receptive vocabulary acquisition is normally linked to passive knowledge and skills like listening and reading while productive vocabulary acquisition is associated to active knowledge and skills like speaking and writing (Laufer & Goldstein, 2004).

3.4.1 Threats to Internal and External Validity

Apart from the participant selection and extraneous variables that may be present in the study as discussed above, some of the possible threats to internal validity of this experimental design are maturation, testing effects and imitation of treatment. Maturation predominantly deals with effects that are caused by time especially in relation the period of time between pre-test, post-test and presentation of treatment. During this experimental period, participants may be affected by the input given during formal learning of English lessons in classrooms, which would be difficult for the researcher to control.

Apart from maturation, testing effects may pose threat to internal validity. A Hawthorne effect may be an inevitable bias, whereby the participants observed may improve in terms of their performance in the classroom when being observed (Landsberger, 1958). Therefore, the researcher used non-reactive measures, which are indirect methods to collect data such as the usage of existing teachers and audio-recordings in this study (Creswell, 2005). Existing teachers would function as the raters and instructors during the treatment. The researcher was not present in the room and only audio recordings during the Listening Comprehension tests and the VKS test would be gathered. With more tests being administered, the participants may score

higher due to the increase exposure of the tested items. In order to overcome this issue, five distractors were added to maintain the motivation and for each test and treatment session, the orders of the item are changed to avoid task familiarity (de la Fuente, 2002).

If participants from the controlled group were to interact with other participants from other groups, imitation of treatment may transpire. The participants from the three groups may perform similarly, which would affect the validity of the findings. It is important that the instructors or teachers inform the participants to not talk anything about the experiment until it is over. Additionally, the tests and treatments are time-based. Therefore, each activity does not exceed 30 minutes to maintain the participants' level of motivation. Each pre-test, post-test and treatment were given to the participants every morning to prevent any forms of fatigue and hunger.

3.5 Procedure

The researcher trained and briefed three existing teachers as instructors and evaluators. This ensures that the treatment and tests run smoothly as the teachers have established classroom rules and the participants view the selected teachers as figures of authority. The teachers were trained 1 week prior to this study. The researcher demonstrated how the experiments and assessments should be conducted for the GPIO, GINW and GINP group. The teachers were also briefed on the purpose and objectives of study. Additionally, the selected teachers possess Bachelor of Education in Teaching English as a Second Language (TESL) and have 6-7 years of teaching experience (the English Language) in primary schools (see Appendix I). To avoid evaluator effect, the instructors for GPIO, GINW and GINP group were assigned different groups on the first and second day. An evaluator effect may affect the results of the study (de la Fuente, 2002). The instructors delivers the instructions prepared by the researcher (see Appendix D1 and Appendix D2). The instructors may modify the instructions to suit the participants' level of proficiency.

3.5.1 Tested Items

The selection of ten concrete nouns is based on the word list to be learnt as stipulated in the Standard Document (Curriculum Development Centre, 2011). The participants have yet to learn the ten target nouns, which will be verified through the results of the pre-test. This increases internal validity (de la Fuente, 2002). As Laufer (1990) pointed out, the noun is the easiest word class to learn followed by verbs, adjectives and adverbs, as it is a direct referral to physical objects and images. Also, previous studies proposed that lexical items that have more than four syllables are harder to learn. Therefore, the selected lexical items do not have more than four syllables (broccoli, carrot, onion, brinjal, cauliflower, pumpkin, celery, cabbage, cucumber and pepper).

3.5.2 Task 1

The purpose of Task 1 was to provide basic exposure of the lexical items, which encompasses the base form and meaning of the lexical items without particular reference to any L2 morphosyntax (Shintani, 2011; de la Fuente, 2002). Task 1 was administered on the first day. Task 1 is a 10-minute listening comprehension task in the form of an information gap activity, with 1 minute for each target lexical item. The task will be conducted in a classroom setting for all groups (GPIO, GINW and GINP). A classroom setting was selected based on the possibility that interaction may not occur in GINW and GINP, as the child may be reluctant to speak if the instructor were to conduct the task individually. Mackey (1998) further suggests that based on empirical evidence, although a child may not interact during the task in a classroom setting, the child benefited in terms of L2 vocabulary acquisition just by listening to the negotiated interaction between his peers and instructor.

Each participant was provided with 15 individual pictures of different vegetables (see Appendix E1). The 15 pictures of vegetables are pictorial representations of the

objects in real life. Previous researchers (Gibson, 1969; Arnheim, 1974; Gibson, 1971) have suggested the positive effects on language learning for young learners due to the nature of similarity of pictures and three-dimensional world. Ten vegetables were the target test items while another five vegetables were included as distractors to maintain motivation (Shintani, 2011; de la Fuente, 2002). Additionally, the participants were provided with a numbered sheet from 1 – 10 (see Appendix E2). Then, the instructor gave the baseline instructions to the participants (see Appendix D1 and F). For the GPIO group, the instructions were delivered at a slower rate. Interactions were not allowed from the participants. However, the participants may ask the instructor to repeat the instructions as many times as needed. The participants were required to paste individual pictures of the vegetables in the box provided on the numbered sheet based on the pre-modified input given. The participants have 1 minute to choose and paste a picture of a vegetable based on the input given. In both GINW and GINP groups, interaction was allowed within the time allotted. The instructions were delivered at a normal speed. The time allotted for each tested item is 1 minute, inclusive of instructions and interaction. One or zero points were awarded based on the accuracy of each target lexical item, with a possible total of 10 points for each task.

Since the task at hand was time-based, the session lasted 10 minutes. The task was repeated and the orders of the test items were altered to avoid task effect. However, the roles of the participant and the instructor will be reversed in the second session for the GINP group. Each participant had to instruct the instructor to locate each target lexical items. The instructors are given 1 minute to listen and negotiate meaning with the participant to paste an individual picture in the box provided on the numbered sheet for each lexical item. Each participant in GINP group was given the opportunity to modify his or her output and produce the tested items in L2. Interactions and questions

on the tested items were allowed from the participants. The purpose of this task was to encourage language production. The total time taken for Task 1 would be 20 minutes.

Both sessions were video-recorded in order to analyze the negotiated interaction that transpired between the instructors and participants. The sample of interaction transcriptions was based on Jefferson's (1974) transcription conventions.

3.5.3 Task 2

The purpose of Task 2 is to increase the exposure of the lexical items to all groups. The session was conducted the following day. This time around, instead of a numbered sheet, each participant was provided with a picture of an empty kitchen with labels on different furniture (see Appendix E3). It is important to note that the participants in all groups had learnt the names of the furniture (fridge, window, floor, table, cabinet and chair) and prepositions used in the instructions (on, in). The procedure is similar to Task 1 including the time allotment and order of the test items presented (see Appendix D2 and F). The task was also repeated.

3.6 Measures

Three testing instruments were administered to the 3 groups (GPIO, GINW and GINP). Firstly, a listening comprehension task, adapted from de la Fuente (2002), to measure vocabulary comprehension. The second and third testing instruments are the two Vocabulary Knowledge Scale (VKS) tests, adapted from Wesche and Paribakht (1996), to measure receptive (RVKS) and productive vocabulary acquisition (PVKS) respectively (see Appendix G).

VKS is classified as a self-reflective five-point assessment scale assessed by the learners for vocabulary acquisition. For the purpose of this study, the VKS test has been modified to a four-point scale. Schmitt (2010) asserts that if the five-point scale has been reduced to four, the test may not be able to gauge the gradual mastery of lexical items beyond word level. However, for the purpose of this study, a four-point scale is

applicable because the research objectives do not intend to measure mastery of lexical items in context. Although Wesche and Paribakht (1996) developed the VKS test to assess adult ELL learners in college, this test is widely available, and has been used in many investigational studies that are primarily concerned with elementary ESL learners (Kütük, 2007; Johannsdottir, 2010; Fisher, Brozo, Frey & Ivey, 2011). The VKS test has been deemed suitable to be modified for the youngest learners (Blachowicz & Fisher, 2006; Stahl & Bravo, 2010). Aside from reporting ESL learners' self-assessment scale on their vocabulary acquisition, the learners would also have to demonstrate their knowledge based on the scales reported. Several studies (de la Fuente, 2002; Stahl & Bravo, 2010) have also revealed that VKS is an effective assessment tool for English language learners and native English speakers alike.

The VKS test for productive acquisition was conducted first to avoid test effects followed by the VKS test for receptive acquisition. Four sets of scores were gathered from one pre-test and three post-tests for receptive and productive vocabulary acquisition. The pre-test was administered before the treatment. The first post-test was administered a day after the treatment to determine the participants' immediate vocabulary retention while the second post-test was administered a week later to determine the participants' delayed vocabulary retention. Finally, the third post-test was administered 3 months later to ascertain the participants' vocabulary retention in the long term. To avoid task effects or familiarity, the order of the items tested was altered. The vocabulary comprehension was administered in the designated groups (GPIO, GINW and GINP) during the experiment where as, the VKS tests for receptive and productive vocabulary acquisitions were administered individually. The processes of the VKS tests were video recorded to ensure consensus in assessment among the three evaluators, who are also the instructors. The total time allotted for the three tests was 30 minutes. The minimum score for each test is 0 and the maximum score is 10.

i. Vocabulary Comprehension

The first listening comprehension task (Task 1) serves as the vocabulary comprehension test. Their vocabulary comprehension is measured based on the scores collected out of 10 while performing the task. The evaluator delivered the instructions provided by the researcher (see Appendix F) to the participants. The participants are required to paste the individual picture in the box provided on the numbered sheet (see Appendix E1, Appendix E2 and Appendix E3). One or zero points were awarded based on the accuracy of each target lexical item, with a possible total of 10 points for each task.

ii. Receptive Vocabulary Acquisition

In order to measure receptive vocabulary acquisition, the receptive VKS test was conducted individually in the evaluator's room. Each participant is provided a receptive four-point vocabulary knowledge scale. The evaluator briefs the four scales to the participant (see Appendix G – Part 1). The VKS test has been deemed suitable to be modified for the youngest learners (Blachowicz & Fisher, 2006; Stahl & Bravo, 2010). In this current study, the participant may point to Scale 1 – if they do not remember having heard the word before, Scale 2 – if they heard the word before, but do not know what it means, Scale 3 – if they heard the word before, and think it means _____ (translated in L1) or Scale 4 – if they know the word and it means _____ (translated in L1 or L2). The evaluator says aloud each tested items twice to the participant. The participant has to listen to each word carefully and select one of the four-point scales that best described their receptive vocabulary acquisition. For the receptive acquisition test, scores are awarded when participants are able to identify and translate the vocabulary spoken by the evaluator in their mother tongue or L2 based on the knowledge of form and meaning (Nation, 2001; Milton, 2009). 0 points are awarded if participants points to Scale 1 or 2. While 1 point will be awarded, if participants points to Scale 3 or 4, provided the participants say aloud the correct translation (or

produced in L2) of the tested item is produced. One minute was allotted for each tested item.

iii. Productive Vocabulary Acquisition

In order to measure receptive vocabulary acquisition, the receptive VKS test was conducted individually in the evaluator's room. Each participant is provided a receptive four-point vocabulary knowledge scale. The evaluator briefs the four scales to the participant (see Appendix G – Part 2). To illustrate, the participant may point to Scale 1 – if they do not know the image, Scale 2 – if they do not remember how to say the image in L2, Scale 3 – if they think it is a _____ (produce in L2) or Scale 4 – if they know it is a _____ (produce in L2). Then, the evaluator shows pictures of each tested item to the participants individually on the evaluator's desk. The participants have to look at the pictures one by one carefully and point to one of the four-point scales that best described their productive vocabulary acquisition. However, their productive acquisition will be measured based on their ability to orally demonstrate each tested item in L2. One minute was allotted for each tested item. The scores are awarded if participants are able to demonstrate and produce the vocabulary in L2 regardless of morphological inaccuracies. To illustrate, if participants are able to provide the whole word irrespective of minimal pronunciation mistakes (eg. '*cauyeefflower*' for 'cauliflower'), scores will be awarded. However, if participants are not able to provide the whole word (eg. '*broco*' for 'broccoli'), no scores will be awarded.

3.7 Validity of Testing Instruments

Validity is an increasingly important area in applied linguistics. The process of validation helps to ensure that the instrument is measuring what it intends to measure by attaining accurate information from the sample. Previous research outlined the three forms of validity; content validity, criterion-related validity and construct validity.

In terms of content validity, a considerable amount of reviewed literature has assisted in the development of the listening comprehension task assessment (Loschky, 1994; de la Fuente, 2002; Shintani, 2011). Apart from that, in the development of the listening comprehension task, a panel of experts, namely two lecturers from Teachers Training Colleges and two lecturers from University of Malaya and University Putra in Malaysia assisted in the validation of content validity. As for the Vocabulary Knowledge Scale, a number of studies has established to usage of scales to measure vocabulary acquisition (Joe, 1994; McNeill, 1996; Scarcella and Zimmerman, 1998). Stahl and Bravo (2010) mentioned that Wesche and Paribakth's VKS was developed based on Dale's (1965) model of vocabulary acquisition stages to refine the scales in terms of content validity.

In order to safeguard against validity and reliability threats, Wesche and Paribakth (1996) conducted a criterion-validity test on the VKS instrument. Their findings revealed that the VKS had a high correlation of .95 the learners' perceived vocabulary knowledge and their demonstrated knowledge in different themed vocabulary. The enhancement of criterion validity allows future researchers or education practitioners to predict future performance of the VKS as the test scores gathered correlates with the variables that are of interest (Youngman, 1979; Thorndike, 1971).

3.8 Reliability

Reliability refers to the accuracy and precision of the measuring instrument and procedure in terms of consistency. The instrument should be able to yield the same results administered with the same procedure and instrument from the same sample (Youngman, 1979; Thorndike, 1971). This would directly increase the reliability of the measuring instrument. In terms of the reliability of the Vocabulary Knowledge Scale used in this study, Wesche and Paribakth (1996) conducted test-retest reliability on ELL

learners in a university of different levels of proficiency. Based on the results of the study, the test-retest correlation is above .8, which indicates that the VKS instrument is high in reliability.

3.9 Data Analysis

The data analysis is twofold. Firstly, the measuring instruments, which are the listening comprehension test and VKS test for both receptive and productive acquisition, are evaluated to enhance its validity and reliability. Although Wesche and Paribakth (1996) conducted criterion-validity and reliability test on the VKS test, the researchers suggested further study in terms of reliability. Therefore, the listening comprehension test and the VKS are subjected to the statistical procedure of test-retest reliability, inter-rater reliability and internal consistency (Davidshofer & Murphy, 2005).

Secondly, to describe and analyze the data gathered. The data set was subjected to the SPSS Version 21 treatment. The data were visually expressed and reported through descriptive statistics followed by inferential statistics as described in detailed by Lay and Khoo (2009). The mean, standard deviation and significance difference of each question are calculated. Various statistical procedures were used to draw meaningful conclusions from the actual data gathered in the actual study. In addition to descriptive analysis tools, a T-test to compare mean, correlational analysis, and ANOVA to compare multiple mean scores were also employed to answer the research questions and to test the hypotheses developed earlier. For research question 1, in order to measure L2 vocabulary comprehension, the significant difference between the mean scores of the first listening comprehension task between two groups (GPIO and GINW + GINP) was analyzed with an independent sample t-test. For research questions 2 and 3, in order to measure receptive and productive vocabulary acquisition, the significant difference between the mean scores of the VKS test for receptive and productive

acquisition of three groups (GPIO, GINW and GINP) was analyzed with a 3 x 3 repeated measures of ANOVA. The repeated measures adheres to the multifactorial design which attempts to analyze the mean scores within-subject design (Pre-test vs. Posttest 1 vs. Posttest 2 vs. Posttest 3) and the mean scores between-subject design (GPIO vs. GINW vs. GINP group). If the ANOVA shows any form of significance, a post hoc Scheffé test will be conducted to determine the group that performed significant better than the other.

Based on the SPSS output, the results will be interpreted and analyzed in the light of relevant literature on the positive effects of negotiated interaction with or without output towards L2 vocabulary acquisition or revealed no significant difference. The investigator triangulation method is directly utilized for confirmation purposes on the SPSS output to enhance the validity of the findings (Guba & Lincoln, 1981; Hartley & Chesworth; 2000; Mills, 2003; Bryman, 2012). Social researchers (Denzin, 1989; Thurmond; 2001) classified investigator triangulation as the process of involving multiple data analysts to check and verify the findings for the same study. For the purpose of this study, two esteemed professionals in statistical measurement have checked the SPSS output for the data analysis and pilot study. The first professional enlisted is Dr. John Michael Linacre, who possesses a doctorate in Educational measurement and was a former Director of MESA Psychometric Laboratory in University of Chicago. The second professional enlisted is Premila Subramaniam, who possesses a Masters degree in Applied Statistics and has served as an Application Consultant in SPSS Malaysia.

3.10 Pilot Study

A pilot study was conducted to 45 pupils from a primary national school in Kuala Langat district. They were selected based on their age, school location, gender and level of proficiency in English. The samples possessed similar traits as the actual

samples and were not included in the final sample of the study. The table below (Table 3.2) summarizes the sample selected for the pilot study. This pilot study intends to test the three aspects of reliability of the listening comprehension test and the VKS tests, namely stability, equivalence and internal consistency. The testing instruments are subjected to multiple statistical procedures of test-retest reliability to assess stability, inter-rater reliability to assess equivalence and split-half method to assess internal consistency (Davidshofer & Murphy, 2005; Trochim, 2006).

Table 3.2 : Pilot Study's Sample Size Based on Stratified Random Sampling for Kuala Langat Division

Groups		Number of Samples	
		Gender	Total
GIPO (Control Group)	Female	8	15
	Male	7	
GINW (Experimental Group)	Female	8	15
	Male	7	
GINP (Experimental Group)	Female	8	15
	Male	7	
Total		45	45
Percentage		100	100

3.10.1 Test-retest Reliability

Test-retest reliability aims to measure the consistency of the results yield by the same participants under the same condition in two different occasions. Three tests (Listening Comprehension Tests, Receptive Vocabulary Knowledge Scale and Productive Vocabulary Knowledge Scale) were administered twice to the selected 45 participants in a lapse of one week between the first and second tests. The presentation and order of the items tested were altered to avoid test effect. The pilot data set would then be subjected to SPSS treatment, whereby the Intraclass Correlation Coefficient (ICC) would be computed. The ICC would be analyzed based on Shrout and Fleiss's (1979) conventions of a two-way mixed single measure, consistency type (3,1).

The Intraclass Correlation Coefficient (ICC) is said to be one of the statistical procedure to draw meaningful conclusion on the significant difference between means for test-retest reliability and inter-rater reliability (Bland & Altman, 1990). Although it resembles Pearson's r , the ICC's r is used to compute the reliability coefficient. The results will be interpreted based on Landis and Koch's (1977) scale as follows; poor to fair (below 0.4), moderate (0.41–0.60), excellent (0.61–0.80), and almost perfect (0.81–1). Landis and Koch's scale is also in line with Lee et al.'s (1989) recommended acceptable scale, which falls greater than .75.

Table 3.3: Descriptive Statistics for Listening Comprehension Test (LCT), Receptive Vocabulary Knowledge Scale (RVKS) and Productive Vocabulary Knowledge Scale (PVKS) for 1-Week Elapse Time

Tests	N	Mean	Difference	SE of Mean	Minimum	Maximum
LCT1	45	2.78	0.20	0.12	2.00	5.00
LCT2	45	2.98		0.12	2.00	5.00
RVKS1	45	2.07	0.02	0.12	0.00	3.00
RVKS2	45	2.09		0.13	0.00	3.00
PVKS1	45	0.42	0.09	0.07	0.00	1.00
PVKS2	45	0.51		0.08	0.00	1.00

Based on the descriptive statistics illustrated in Table 3.3, a test-retest was administered to 45 participants in a lapse of one-week time between each test. The mean scores for LCT1, RVKS1 and PVKS1 are 2.78, 2.07 and 0.42 with a standard error of mean of 0.12, 0.12 and 0.07 respectively. The mean scores for LCT2, RVKS2 and PVKS2 are 2.98, 2.09 and 0.51 with a standard error of mean of 0.12, 0.13 and 0.08 respectively. The differences of mean scores between LCT1 and LCT2 is 0.20, RVKS1 and RVKS2 is 0.02, while the differences of mean scores between PKVS1 and PVKS2 is 0.09. The positive value of the mean differences indicates that the participants scored higher in their second test in comparison to their first test.

The test-retest reliability obtained from the ICC analysis is shown in Table 3.4. The ICC for LCT records a .809, which indicates acceptable and excellent agreement. Additionally, the Cronbach's Alpha for LCT is .0895, which also indicates good internal consistency between the two tests administered (George & Mallery, 2003). The ICC for RVKS is .778, which indicates acceptable and excellent agreement. The Cronbach's Alpha for RVKS is .0875, which also indicates good internal consistency between the two tests administered. The ICC for PVKS is .836, which indicates acceptable and excellent agreement. The Cronbach's Alpha for PVKS is .0911, which also indicates excellent internal consistency between the two tests administered.

Table 3.4: Inter-rater Intraclass Correlation Coefficients and Cronbach's Alpha for Two Tests

Tests	Intraclass Correlation Coefficients			Reliability Statistics	
	ICC	95% Confidence Interval		Cronbach's Alpha	N of Items
		Lower Bound	Upper Bound		
LCT	.809	.678	.891	.895	2
RVKS	.778	.630	.872	.875	2
PVKS	.836	.720	.906	.911	2

3.10.2 Inter-rater Reliability

Inter-rater reliability is primarily concerned with the level of agreement in the scores awarded by different raters, to the same participants of the same test (Saal, Downey & Lahey, 1980). Shuttleworth (2009) asserts that it is necessary to conduct an inter-rater reliability upon assessments that requires subjective ratings from trained or untrained raters. The scores awarded for RVKS and PVKS test are subjected to the rater's interpretation of the translated tested items and the pronunciation of the tested items in English. Therefore, an inter-rater reliability is essential to establish an acceptable level of agreement among the raters. The raters are selected based on their qualifications. They possess a Bachelor's degree in TESL and have been servicing in primary national schools for more than 5 years as English teachers. The scores between

the two raters would then be subjected to SPSS treatment, whereby the ICC would be computed. The ICC, similar to the test-retest reliability, is analyzed based on Shrout and Fleiss's (1979) conventions of a two-way mixed single measure, consistency type (3,1). Rater 1 is labeled as R1 and Rater 2 is labeled as R2.

Table 3.5: Descriptive Statistics for Receptive Vocabulary Knowledge Scale (RVKS) and Productive Vocabulary Knowledge Scale (PVKS) for Two Raters

Tests	N	Mean	Difference	SE of Mean	Minimum	Maximum
RVKSR1	45	2.07	0.00	0.12	0.00	3.00
RVKSR2	45	2.07		0.12	0.00	3.00
PVKSR1	45	0.42	0.06	0.07	0.00	1.00
PVKSR2	45	0.36		0.07	0.00	1.00

Based on the descriptive statistics illustrated in Table 3.5, the RVKS and PVKS tests were administered to 45 participants and were rated by two raters at the same time. The mean scores for RVKSR1 and PVKSR1 are 2.07 and 0.42 with a standard error of mean of 0.12 and 0.07 respectively. The mean scores for RVKSR2 AND PVKSR2 are 2.07 and 0.36 with a standard error of mean of 0.12 and 0.07. The mean difference between RVKSR1 and RVKSR2 is 0.00, which indicates that the raters rated the participants almost the same. The mean difference between PVKSR1 and PVKSR2 is 0.06, which indicates that Rater 1 rated the participants higher than Rater 2.

Table 3.6: Inter-rater Intraclass Correlation Coefficients and Cronbach's Alpha for Two Raters

Tests	ICC	Intraclass Correlation Coefficients		Reliability Statistics	
		95% Confidence Interval		Cronbach's Alpha	N of Items
		Lower Bound	Upper Bound		
RVKS	.963	.933	.979	.981	2
PVKS	.868	.773	.926	.930	2

It can be seen from the data in Table 3.6 that, the ICC for RVKS records a high .963 with a Cronbach's Alpha value of .981, which indicates high level of reliability and an almost perfect agreement between the two raters in terms of the scores awarded. On the other hand, the ICC for PVKS is .868 with a Cronbach's Alpha value of .930, which indicates a high level of reliability and excellent level of agreement between the two raters for the scores awarded.

3.10.3 Split-half Reliability

Internal consistency reliability is primarily concerned with the consistency of performances on one group of individual across items on a single measure (Cozby, 2001; Streiner, 2003). One of the methods to establish internal consistency is through split-half reliability, whereby the test administered is split into halves based on the evens-odds approach.

$$\alpha = \frac{K}{K-1} \left[1 - \frac{\sum_{i=1}^N p_i q_i}{\sigma_x^2} \right]$$

α = Alpha statistic

K = Number of items in the test

p = Probability of i

$q = 1 - p$

σ = Standard deviation of the distribution

N = Number in population

Figure 3.2: The Formula for KR-20 for a Test cited in Kuder and Richardson, 1937

Each halved scores was then compared and calculated based on the Kuder-Richardson Formula 20 as seen in Figure 3.2. Kuder-Richardson Formula 20 (KR-20) is

often used to calculate the reliability coefficient and extensively utilized in social research methods (Kuder & Richardson, 1937; Rea, 1997). The KR-20 is particularly useful to measure test items that are of dichotomous choices, in this case, 1 point is awarded for the correct answer and 0 point is awarded for incorrect answers.

The results will be interpreted based on Haladyna (1999) scale as follows; unreliable (.0), not reliable (<.5), very reliable (>.8) and completely reliable (1.0). The KR-20 value for LCT, RVKS and PVKS is .85, .86 and .91 respectively. These values indicate high level of reliability in internal consistency.

3.11 Conclusion

This chapter serves to present the analysis of the measuring instrument and procedure before the actual study is being carried out. SPSS are performed on the VKS test to extend its reliability based on the findings of the pilot data set. There is a need for pilot study and analysis to ensure that the instruments yield accurate results and the treatment session to run smoothly by rectifying any shortcomings of the procedures.

CHAPTER 4: FINDINGS

4.1 Introduction

The purpose of this study is to investigate the effects of different oral interactions on new L2 vocabulary acquisition among Primary ESL learners in Kuala Langat, Malaysia. Studies by Pica, Doughty and Young (1987), de la Fuente (2002), Ng and Sheila Marie Sappathy (2011) revealed that negotiated interaction and input is effective in enhancing L2 vocabulary comprehension for adults and young learners alike. Therefore, it was firstly hypothesized that learners who are exposed to pre-modified input with negotiated interaction will attain higher level of vocabulary comprehension than those learners who are exposed to pre-modified input only. Studies by Ellis & He (1999) and de la Fuente (2002) have shown that negotiated input and output production have been proven to be beneficial for receptive and productive acquisition. Based on these studies, it was hypothesized that learners who are exposed to pre-modified input with negotiated interaction with or without output production will attain higher level of receptive and productive vocabulary acquisition than those learners who are exposed to pre-modified input only. In the attempt to approve or disprove the hypotheses, three corresponding research questions and hypotheses has been formulated to ascertain the relative effects of different oral interactions in the primary ESL learners' L2 vocabulary comprehension and acquisition.

In this present study, the quantitative data were gathered from three testing instruments as illustrated below. The three independent variables determined the type of intervention subjected upon each group (GPIO – pre-modified input, GINW – negotiated input and GINP – negotiated input with output production). The three dependent variables that are under investigation were:

- i. Vocabulary comprehension: This was measured by the listening comprehension task administered in the first round of the treatment. The

participants are required to match the pictures to their corresponding numbers based on the instructions delivered. The maximum score was 10.

- ii. Receptive vocabulary acquisition: This was measured with the Vocabulary Knowledge Scale on four separate occasions (Pretest, Posttest 1, Posttest 2 and Posttest 3). The participants are required to rate themselves under the scale that best represent them and are to produce the target word in their mother tongue or point to the correct picture. The maximum score was 10.
- iii. Productive Vocabulary Acquisition: Similar to receptive vocabulary acquisition, the scores are obtained from four separate occasions (Pretest, Posttest 1, Posttest 2 and Posttest 3). The participants are required to rate themselves under the scale that best represent them and are to produce the target word in the L2 with minimal pronunciation mistakes. The maximum score was 10.

Therefore, five sets of scores were gathered from the single listening comprehension test, pre-tests and a series of three posttests administered a day after, a week after and three months later from the Vocabulary Knowledge Scale. Before subjecting the data to any forms of inferential statistics for hypothesis testing, the data needs to fulfill normality assumption and homogeneity of variance assumption. The normality of variances was subjected to the Kolmogorov-Smirnov and Shapiro-Wilk normality tests and the homogeneity of variances were subjected to Levene's test. When the assumptions are fulfilled, the data were subjected to the appropriate inferential analysis.

This chapter is divided into four main sections. The first section of the results comprises of the pre-test analysis of the VKS for the three groups; the control group (GPIO), the experimental groups (GINW and GINP) with Analysis of Variance (ANOVA) to establish the equality in terms of the proficiency level in English. This will be followed by the three other sections, which attempted to answer to the three

hypotheses with their corresponding research questions. In order to measure vocabulary comprehension, the mean test scores between the GPIO and GINW with GINP were compared with independent sample t-test. To address hypothesis 2 and 3, 3 x 3 repeated measures of ANOVA was employed to compare the mean scores between-groups and within-group factor. For the between-groups factor, the mean scores of the pretest and the immediate posttest were analyzed. This is to check on the effectiveness of negotiated input and output in terms of the immediate retention of vocabulary between groups. This procedure is followed by the within-group factor measurement; the mean scores for the pretest and three posttests were compared to garner significant difference within each group through ANOVA. This is to establish the effectiveness of negotiated input and output production in relation to time factor. Finally, when the 3 x 3 repeated measures of ANOVA showed significant difference in terms of the mean scores between-groups and within group, post hoc Scheffe tests were used to confirm the hypotheses. In order to interpret the results, the significance level is determined at 0.5 level of significance (alpha, $\alpha = .05$) or at a 95% of confidence interval. Scores are read to reach statistical significance if the confidence levels are non-overlapping. If the results have overlapping confidence intervals, they may not be necessarily reaching statistical significance.

4.2 Analysis of Pretest

The results of the pretest aimed at establishing the assumption of equality of variance are presented in Tables 4.1 and 4.2 below. The null hypothesis to be tested ($H_0: \mu_E = \mu_C$) states that the RVKS and PKVS pretest mean scores of the experimental groups are equal to the pretest mean scores of the control group. Conversely, the alternative hypothesis ($H_1: \mu_E \neq \mu_C$) states that the pretests RVKS and PVKS mean scores of the experimental groups are not equal to the pretests mean scores of control group. The significance level alpha is specified at .05. Table 4.1 presents the descriptive

statistics of the pretest scores obtained from the GPIO (control group), GINW and GINP (experimental groups) of the RVKS and PVKS. For RVKS, the participants from the GPIO ($M = .93$, $SE = .799$) scored slightly higher than the experimental groups (GINW $M = .80$, $SE = .775$, GINP $M = .80$, $SE = .775$). Conversely for PKVS, both the experimental groups (GINW $M = .80$, $SE = .676$, GINP $M = .80$, $SE = .862$) scored slightly higher than the control group (GPIO $M = .53$, $SE = .516$).

Table 4.1 Descriptive Summary of Pretest for RVKS and PVKS

Tests	Groups	N	Mean	Std. Deviation	Std. Error Mean
RVKS	GPIO	15	.93	.799	.206
	GINW	15	.80	.775	.200
	GINP	15	.80	.775	.200
PVKS	GPIO	15	.53	.516	.133
	GINW	15	.80	.676	.175
	GINP	15	.80	.862	.223

Table 4.2 Test of Homogeneity of Variances

	Levene Statistic	df1	df2	Sig.
RVKS	.009	2	42	.991
PVKS	2.726	2	42	.077

A Shapiro-Wilk's test ($p > .05$) (Shapiro & Wilk, 1965; Nornadiah Razali & Wah, 2011) and a visual inspection of their histograms, normal Q-Q plots and box plots showed that the pretest scores were approximately normally distributed for the three groups with a skewness of .128 ($SE = .580$) and a kurtosis of -1.348 ($SE = 1.121$) for GPIO, a skewness of .383 ($SE = .580$) and a kurtosis of -1.117 ($SE = 1.121$) for GINW and a skewness of .383 ($SE = .580$) and a kurtosis of -1.117 ($SE = 1.121$) for GINP

(George & Mallery, 2003; Cramer & Howitt, 2004; Doane & Seward, 2011). Based on Table 4.2, the F value of the RVKS for Levene's test is .009 with a Sig. (p) value of .991 while the F value of the PVKS for Levene's test is 2.726 with a Sig. (p) value of .077. Therefore, the null hypothesis (no difference) for the assumption of homogeneity of variance is retained and concludes that there is no significant difference between the three group's variances because the Sig. value is greater than the alpha of .05 ($p > .05$). A Levene's test verified the equality of variances in the samples (homogeneity of variance) ($p > .05$) (Martin & Bridgmon, 2012) where the pretest scores between three groups are deemed to be similar. The results displayed in Tables 4.1 and 4.2 indicate that the participants across the groups possessed similar level of English proficiency and knowledge of vocabulary. Therefore, the three groups were comparable prior to treatment.

4.3 Effects of Negotiated Interaction on Comprehension

Research Question 1: What are the relative effects of pre-modified input and negotiated interaction on L2 vocabulary comprehension?

In order to ascertain the significant difference between the Listening Comprehension Task (LCT) test scores of the group with pre-modified input (GPIO) and the group with negotiated input (GINW and GINP), an independent samples t-test was employed to analyze the mean scores obtained by the control and experimental groups. The GINW and GINP test scores are combined because both groups have a similar variable, which is negotiated interaction. Both GINW and GINP group performed the same interactive task during the first round of the listening comprehension task. Tables 4.3 and 4.4 display the mean scores, standard deviations, (SD), and t value for the comprehension task. The significance level is determined at 0.5 level of significance (alpha, $\alpha = .05$) or at a 95% of confidence interval. The null hypothesis tested (H_0) for research questions 1 states that learners who are exposed to

pre-modified input with negotiated interaction will attain equal level of vocabulary comprehension to the learners who are exposed to pre-modified input only. The alternative hypothesis tested (H_1) states that learners who are exposed to pre-modified input with negotiated interaction will attain higher level of vocabulary comprehension to the learners who are exposed to pre-modified input only.

Table 4.3 Descriptive Summary of LCT

	Group	N	Mean	Std. Deviation	Std. Error Mean
LCT	GPIO	15	4.00	.845	.218
	GINW + GINP	30	7.70	.952	.174

Table 4.4 Independent Samples Test for LCT

Levene's Test for Equality of Variances		t-test for Equality of Means						
F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
							Lower	Upper
.833	.366	-12.735	43	.000	-3.700	.291	-4.286	-3.114

Based on Tables 4.3, the group with negotiated input (GINW and GINP $M = 7.70$, $SE = .952$) scored higher than the groups with pre-modified input only (GPIO $M = 4.00$, $SE = .845$). Based on the results of the independent samples t-test, $t(43) = -12.735$, $p = .000$, 95% CI [-4.28, -3.11], since the significant value was lesser than alpha at 0.5 level of significance, there was sufficient evidence to reject null hypothesis. It can be concluded that negotiated input had a significant effect on the primary ESL learners' L2 vocabulary comprehension.

4.4 Effects of Negotiated Interaction on Receptive Vocabulary Acquisition

Research Question 2: What are the relative effects of pre-modified input, negotiated interaction with and without output production on L2 receptive vocabulary acquisition? If there is an effect, how does it affect L2 vocabulary acquisition in terms of retention (time)?

In order to ascertain the significant difference of the Receptive Vocabulary Acquisition Scale (RVKS) test scores from the three groups of different types of negotiation (GPIO, GINW and GINP), a 3x3 mixed design ANOVA was employed to analyze the mean scores obtained by the control and experimental groups. The ANOVA will provide the between-subjects test effects, which was the type of negotiation (GPIO vs. GINW VS. GINP) and the within-subject test effects, which was the time factor (Test 1 vs. Test 2 vs. Test 3).

Table 4.5 displays the mean scores (*M*) and standard deviations (*SD*) for the three posttests. The significance level is determined at 0.5 level of significance (alpha, $\alpha = .05$) or at a 95% of confidence interval. The null hypothesis tested (H_0) for research questions 2 states that there is no significant difference in the receptive vocabulary acquisition scores between learners who are exposed to pre-modified input with negotiated interaction with or without output production and learners who are exposed to pre-modified input only. Conversely, the alternative hypothesis tested (H_1) states that learners who are exposed to pre-modified input with negotiated interaction with or without output production will attain higher level of receptive vocabulary acquisition than those learners who are exposed to pre-modified input only.

Table 4.5 Means (M) and Standard Deviations (SD) for RVKS

Group	<i>n</i>	Posttest 1		Posttest 2		Posttest 3	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
GPIO	15	3.87	.743	3.87	1.187	4.13	1.060
GINW	15	7.13	.743	7.27	.799	7.13	.834
GINP	15	7.60	.828	7.73	.799	7.80	.775

Table 4.6 ANOVA for Receptive Acquisition

Source	SS	df	MS	F	Sig.
Corrected Model	372.770	8	46.596	60.820	.000
Intercept	5326.696	1	5326.696	6952.663	.000
Time	.548	2	.274	.358	.700
T. of Negotiation	371.570	2	185.785	242.496	.000
Time x T. of Negotiation	.652	4	.163	.213	.931
Error	96.533	126	.766		
Total	5796.000	135			
Corrected Total	469.304	134			

Based on the descriptive statistics in Table 4.5, overall, the GINW (Posttest 1 $M = 7.13$, $SD = .743$, Posttest 2 $M = 7.27$, $SD = .799$, Posttest 3 $M = 7.13$, $SD = .834$) and GINP (Posttest 1 $M = 7.60$, $SD = .828$, Posttest 2 $M = 7.73$, $SD = .799$, Posttest 3 $M = 7.80$, $SD = .775$) groups scored higher than the GPIO group (Posttest 1 $M = 3.87$, $SD = .743$, Posttest 2 $M = 3.87$, $SD = 1.187$, Posttest 3 $M = 4.13$, $SD = 1.060$) in all three posttests. The output of the ANOVA in Table 4.6 reported significance between groups in terms of types of negotiation, F value of 242.496, $p = .000$. Based on the ANOVA results, since the significant value was lesser than alpha at 0.5 level of significance, there was sufficient evidence to reject null hypothesis. It can be concluded that the type of negotiation had significant effects on the learners' L2 receptive vocabulary acquisition. However, the output of the ANOVA reported no significant difference in

terms of the time factor, F value of .358, $p = .700$ and the interaction between types of negotiation and time, F value of .213, $p = .931$. Based on the ANOVA results, since the significant value were greater than alpha at 0.5 level of significance, there was no sufficient evidence to reject null hypothesis. It can be concluded that time and interactions between types of negotiation and time have no significant effects on the learners' L2 receptive vocabulary acquisition.

Table 4.7 Scheffé Post Hoc Comparisons for Receptive Acquisition by Group

(I) T. of Negotiation	(J) T. of Negotiation	Mean Difference (I – J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
GPIO	GINW	-3.22*	.185	.000	-3.68	-2.77
	GINP	-3.76	.185	.000	-4.21	-3.30
GINW	GPIO	3.22*	.185	.000	2.77	3.68
	GINP	-.53	.185	.015	-.99	-.08
GINP	GPIO	3.76*	.185	.000	3.30	4.21
	GINW	.53	.185	.015	.08	.99

*. The mean difference is significant at the 0.001 level

In order to determine the significance difference between each groups, a Scheffé post hoc comparison for receptive acquisition was conducted. The significance level is determined at 0.001 level of significance (alpha, $\alpha = .001$). Based on the results as illustrated in Table 4.7, there are significant differences between two combinations of types of negotiation. The overall significant difference between the GPIO group and the GINW group is (3.96 vs. 7.18) $p = .000$, the GPIO group and the GINP group is (3.96 vs. 7.71) $p = .000$, while the GINW group and the GINP group is (7.18 vs. 7.71) $p = .015$. In other words, both negotiated input with and without output group (GINW and GINP) outperformed the group with pre-modified input. However, there was no significant difference between the group with negotiated input with output production (GINP) and the group with negotiated input (GINW). The post hoc comparison was not

used for time, as there was no effects of time on receptive vocabulary retention were found. Briefly put, none of the conditions of the experiment was superior to another in allowing learners to retain the target words receptively.

4.5 Effects of Negotiated Interaction on Productive Vocabulary Acquisition

Research Question 3: What are the relative effects of pre-modified input, negotiated interaction with and without output production on L2 productive vocabulary acquisition? If there is an effect, how does it affect L2 vocabulary acquisition in terms of retention (time)?

To measure the significant difference of the Productive Vocabulary Acquisition Scale (PVKS) test scores from the three groups of different types of negotiation (GPIO, GINW and GINP), similar statistical procedure used for Research Question 2 is used in Research Question 3.

Table 4.8 shows the mean scores (*M*) and standard deviations (*SD*) for the three posttests. The significance level is determined at 0.5 level of significance (alpha, $\alpha = .05$) or at a 95% of confidence interval. The null hypothesis tested (H_0) for research questions 3 states that there is no significant difference in the productive vocabulary acquisition scores between learners who are exposed to pre-modified input with negotiated interaction with or without output production and those learners who are exposed to pre-modified input only. The alternative hypothesis tested (H_1) states that learners who are exposed to pre-modified input with negotiated interaction with or without output production will attain higher level of productive vocabulary acquisition than those learners who are exposed to pre-modified input only.

Table 4.8 Means (M) and Standard Deviations (SD) for PVKS

Group	<i>n</i>	Posttest 1		Posttest 2		Posttest 3	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
GPIO	15	2.27	.458	2.40	.507	2.67	.488
GINW	15	2.47	.516	2.73	.594	3.00	.655
GINP	15	5.20	.561	5.27	.594	8.07	.594

Table 4.9 ANOVA for Productive Acquisition

Source	SS	df	MS	F	Sig.
Corrected Model	471.970	8	58.996	191.586	.000
Intercept	1934.230	1	1934.230	6281.261	.000
Time	42.948	2	21.474	69.735	.000
T. of Negotiation	388.281	2	194.141	630.457	.000
Time x T. of Negotiation	40.741	4	10.185	33.076	.000
Error	38.800	126	.308		
Total	2445.000	135			
Corrected Total	510.770	134			

Based on the descriptive statistics in Table 4.8, overall, the GINP group (Posttest 1 $M = 5.20$, $SD = .561$, Posttest 2 $M = 5.27$, $SD = .594$, Posttest 3 $M = 8.07$, $SD = .594$) scored higher than the GPIO (Posttest 1 $M = 2.27$, $SD = .458$, Posttest 2 $M = 2.40$, $SD = .507$, Posttest 3 $M = 2.67$, $SD = .488$) and GINW groups (Posttest 1 $M = 2.47$, $SD = .516$, Posttest 2 $M = 2.73$, $SD = .594$, Posttest 3 $M = 3.00$, $SD = .655$) in all three posttests. The output of the ANOVA in Table 4.9 reported significance between groups in terms of types of time factor, F value of 69.735, $p = .000$ and the types of negotiation, F value of 634.457, $p = .000$. Apart from that, the output of the ANOVA also reported significant effects in terms the interaction between types of negotiation and time, F value of 33.076, $p = .000$. Based on the ANOVA results, since the significant value was lesser than alpha at 0.5 level of significance, there was sufficient evidence to reject null hypothesis. It can be concluded that the time taken for all the posttests and the type of

negotiation had significant effects on the learners' L2 productive vocabulary acquisition.

Table 4.10 ANOVA for Productive Vocabulary Acquisition by Test

Test/Source	df	SS	MS	F	Sig.
Test 1					
T. of negotiation	2	80.578	40.289	152.904	.000
Residual	42	11.067	.263		
Test 2					
T. of negotiation	2	73.733	36.867	114.980	.000
Residual	42	13.467	.321		
Test 3					
T. of negotiation	2	274.711	137.356	404.364	.000
Residual	42	14.267	.340		

In order to determine the interaction between time and type of negotiation, the mean scores for the three groups of the three posttests were analyzed individually with ANOVA. Based on the output as displayed in Table 4.10, there were significant differences for the type of negotiation on Test 1, $F(2,42) = 152.904, p = .000$, Test 2, $F(2,42) = 114.980, p = .000$ and Test 3, $F(2, 42) = 404.364, p = .000$.

To measure the significant differences between each group, a Scheffé post hoc comparisons for receptive acquisition was conducted as illustrated in Table 4.11. The significance level is determined at 0.001 level of significance (alpha, $\alpha = .001$) or 99% confidence interval. Based on the results provided in Table 4.11, there were no significant differences between the GPIO and the GINW group in the three tests (Test 1 2.27 vs. 2.47, $p = .570$; Test 2 2.40 vs. 2.73, $p = .283$; Test 3 2.67 vs. 3.00, $p = .304$). In contrast, the post hoc comparisons revealed significant differences between the GPIO and GINP group in the three tests (Test 1 2.27 vs. 5.20, $p = .000$; Test 2 2.73 vs. 5.27, $p = .000$; Test 3 3.13 vs. 8.07, $p = .000$). Similarly, there were significant differences

between the GINW and GINP group in all the tests (Test 1 2.47 vs. 5.20, $p = .000$; Test 2 2.73 vs. 5.27, $p = .000$; Test 3 3.13 vs. 8.07, $p = .000$). This indicates that the GINP group managed to produce more words significantly than the GPIO and GINW group for short-term and long-term retention.

Table 4.11 Scheffé Post Hoc Comparisons for Receptive Acquisition by Test

Dependent Variable	(I) T. of negotiation	(J) T. of negotiation	Mean Difference (I-J)	Std. Error	Sig.	99.9% Confidence Interval	
						Lower Bound	Upper Bound
Test 1	GPIO	GINW	-.200	.187	.570	-.96	.56
		GINP	-2.933*	.187	.000	-3.69	-2.18
	GINW	GPIO	.200	.187	.570	-.56	.96
		GINP	-2.733*	.187	.000	-3.49	-1.98
	GINP	GPIO	2.933*	.187	.000	2.18	3.69
		GINW	2.733*	.187	.000	1.98	3.49
Test 2	GPIO	GINW	-.333	.207	.283	-1.17	.50
		GINP	-2.867*	.207	.000	-3.70	-2.03
	GINW	GPIO	.333	.207	.283	-.50	1.17
		GINP	-2.533*	.207	.000	-3.37	-1.70
	GINP	GPIO	2.867*	.207	.000	2.03	3.70
		GINW	2.533*	.207	.000	1.70	3.37
Test 3	GPIO	GINW	.133	.230	.304	-.80	1.06
		GINP	-4.933*	.230	.000	-5.86	-4.00
	GINW	GPIO	-.133	.230	.304	-1.06	.80
		GINP	-5.067*	.230	.000	-6.00	-4.14
	GINP	GPIO	4.933*	.230	.000	4.00	5.86
		GINW	5.067*	.230	.000	4.14	6.00

* The mean difference is significant at the 0.001 level.

To find out the significant difference of time on L2 productive vocabulary acquisition, the mean scores for the three groups of the three posttests were subjected to repeated measures of ANOVA with one-within-subject (Test 1 vs. Test 2 vs. Test 3). The significance level is determined at 0.05 level of significance ($\alpha = .001$) or 95% confidence interval. The ANOVA results presented in Table 4.12 revealed that there was no significant difference for the two groups (GPIO $F = 2.649$, $p = .083$;

GINW $F = 3.055$, $p = .058$). Conversely, there was a significant effect of time for the GINP group ($F = 118.215$, $p = .000$).

Table 4.12 ANOVA for Productive Acquisition by Group

		SS	df	MS	F	Sig.
GPIO	Between Groups	1.244	2	.622	2.649	.083
	Within Groups	9.867	42	.235		
	Total	11.111	44			
GINW	Between Groups	2.133	2	1.067	3.055	.058
	Within Groups	14.667	42	.349		
	Total	16.800	44			
GINP	Between Groups	80.311	2	40.156	118.215	.000
	Within Groups	14.267	42	.340		
	Total	94.578	44			

Based on Table 4.13, the mean comparisons for the GINP group showed no significant differences between Test 1 and Test 2 (5.20 vs. 5.27, $p = .952$). On the other hand, there were significant differences in two combinations of groups; Test 1 and Test 3 (5.20 vs. 8.07, $p = .000$) and Test 2 and Test 3 (5.27 vs. 8.07, $p = .000$). Based on the post hoc results, since the significant value was lesser than alpha at 0.5 level of significance, there was sufficient evidence to reject null hypothesis. This indicates that the participants from the GINP group are able to produce more words significantly for the third test.

Table 4.13 Mean Comparisons for Productive Acquisition: GINP Group

(I) Time	(J) Time	Mean Difference (I – J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Test 1	Test 2	-.067	.213	.952	-.61	.47
	Test 3	-2.867*	.213	.000	-3.34	-2.33
Test 2	Test 1	.067	.213	.952	-.47	.61
	Test 3	-2.800*	.213	.000	-3.34	-2.26
Test 3	Test 1	2.867*	.213	.000	2.33	3.41
	Test 2	2.800*	.213	.000	2.26	3.34

*. The mean difference is significant at the 0.05 level

4.6 Conclusion

This chapter has presented the results of the tests scores through descriptive and inferential statistics. The report provided answers and analysis for each research question and hypothesis, which is cautiously interpreted in the light of relevant literature. The investigator triangulation method is also directly utilized for confirmation purposes on the SPSS output to enhance the validity of the findings (Guba & Lincoln, 1981; Hartley & Chesworth; 2000; Mills, 2003; Bryman, 2004).

CHAPTER FIVE: DISCUSSION

5.1 Introduction

This chapter attempts to summarize the findings of the study and provides suggestions and recommendations for future researches. The first portion of the chapter will discuss the hypotheses of the research through the lenses of various theoretical points of view. The second portion of the chapter includes a summary of the significant results, which will be thoroughly described. This will be accompanied by the explanation of theoretical and pedagogical implications of negotiated interaction for primary ESL learners. The final part of the chapter will provide multifarious recommendations on different methodology and angles for further analysis.

5.2 Discussion of Research Results

5.2.1 Hypothesis 1

The first hypothesis predicted that learners who are exposed to pre-modified input with negotiated interaction would attain higher level of L2 vocabulary comprehension than those learners who are exposed to pre-modified input only is confirmed by the findings of the study. The results suggest that the groups with negotiated input attained higher level of vocabulary comprehension than the pre-modified input group. For the negotiated input groups, the participants are able to obtain additional information of the target lexical items through negotiation of meaning, which may have aided them with particular reference to vocabulary comprehension. The findings of this study, which corroborate the findings of a great deal of the previous work in this field (Ellis et al., 1994; Loschky, 1994, de la Fuente, 2002). In contrast, a study by Ellis and He (1999) found that there were no significant difference between negotiated input and pre-modified input on L2 vocabulary comprehension. An important departure of this study from Ellis and He (1999) is the time taken to complete the task as argued by de la Fuente (2002). In this present study, the participants are

given 1 minute for each lexical item, a total of 10 minutes for each task. The time factor is one of the controlled variables across groups. This ensures that each participant in the three groups had the same amount of time to negotiate for meaning of each target words. Conversely, participants took 45 minutes to complete a task in Ellis and He's study (1999), which may affect the reliability of the findings (Youngman, 1979; Thorndike, 1971).

The findings of this study also suggest that all participants are able to achieve reasonable level of vocabulary comprehension in two input conditions; pre-modified input and negotiated input. This may be credited to the types of task used in the experiment, which is a listen and do tasks in forms of information gap (see Appendix D1, Appendix D2, Appendix E1, Appendix E2 and Appendix E3). Apart from that, all groups are provided with similar input, which would assist the participants in inferring the meaning of the words. The descriptions (see Appendix F) of the lexical item in the input given may have aided the participants from the GPIO group (pre-modified input) to identify and locate the lexical items. For an example, the adjectives 'long' and 'orange' aided in describing and identifying a carrot. As for the GINW and GINP group (negotiated input), the oral provision of the required information from the instructors would have guided the participants such as 'a small tree' to describe the physical appearance of a broccoli (see Interaction Excerpt - Task 2). Briefly put, contextual information provided verbally as comprehensible input promotes L2 vocabulary comprehension in this situation.

In order to ascertain the specific methods used in negotiation, which contributes to the significant difference, examples are drawn from the observation during the treatment. Firstly, participants from the GINW and GINP group, who are given the opportunity to negotiate, are able to control the type and flow of the input received. While controlling and managing the input received, participants are engaged in the

process of achieving input comprehensibility (Krashen, 1985). Subconsciously, the participants would have to produce and repeat the lexical items when there is a need to garner the required information as illustrated below.

(1) Interaction (Task 2)

Instructor: Put the broccoli on the cabinet.

Participant: What colour is broccoli?

Instructor: It is green in colour.

Participant: Is broccoli look like small tree?

Instructor: Yes, it is green and it looks like a small tree.

Similar to studies by Ellis and He (1999) and de la Fuente (2002), apart from the unintentional repetition the negotiated input group are able to yield specific types of information need to enhance vocabulary comprehension. Based on the interaction above, the participants are able to infer the meaning of the target words based on the information asked and received such as colour (green) and physical shape (small tree) of the each lexical item. This additional information resulted from the interaction remained unavailable for the GPIO group. The findings of this study suggests that negotiated input promotes better L2 vocabulary comprehension than understanding lexical items based on pre-modified input alone.

5.2.2 Hypothesis 2

The second hypothesis predicted that learners who are exposed to pre-modified input with negotiated interaction with or without output production will attain higher level of receptive vocabulary acquisition than those learners who are exposed to pre-modified input only is confirmed by the findings of the study. All input conditions showed reasonable level of receptive vocabulary acquisition; pre-modified input, negotiated input without output and negotiated input with output production. It also appears that time and interactions between types of negotiation and time have no significant effects on the learners' L2 receptive vocabulary acquisition. However, the results of the study suggest clear evidence that the GINW and the GINP gained greater

advantage than the GPIO group. The findings of the current study are consistent with those of Ellis and He (1999), in which there is a significant effect of negotiation on L2 receptive vocabulary acquisition in comparison to pre-modified input only group. This contrasted studies by Loschky (1994) and de la Fuente (2002), who found no significant difference between the group with negotiated interaction and the group with pre-modified input only. There was also a significance difference between the GINW and GINP group where the GINP performed significantly better than the GINW group. This result supported previous studies by Ellis and He (1999) and de la Fuente (2002), where the output group's performance are proven superior to the other input conditions in L2 receptive vocabulary acquisition and retention.

In order to ascertain the aspects of GINP that contribute to the differential effects over GINW and GPIO, the process will be interpreted in the light of Swain's (1995) output hypothesis. Based on the treatment procedures, the GINP group receives additional task, in which the participants were pushed to produce the target words in L2. After the first round of the task, they have to assume reverse roles and instruct the instructors to place the target lexical items by producing the same input.

Swain (1995) argues that the processes involved in comprehending a message (decode) and determining the choice of language used to produce (code-breaking) are different (Cook, 1991). When pushed to produce, the participants are obligated to focus on the linguistic features and syntax to convey the instructions meaningfully to the instructors. This process provides participants with various rooms to experiment with linguistic forms. Essentially, the participants engaged in deeper input processing level and retained control over his or her learning, which in turns facilitates L2 productive vocabulary acquisition. In the examination of negotiation on L2 grammatical acquisition, encouraging results have been published by recent studies to push forth the importance of output production for L2 acquisition (Mackey, 1995; Mackey & Philp,

1998; Van den Branden, 1997).

One of the three functions of output hypothesis postulated by Swain (1995) is raising consciousness among language learners to notice the gap. While producing in the targeted language, participants from the GINP group are able to pay attention notice the gap between their current level of linguistic knowledge and what the instructors produce (Schmidt, 1990, 1993, 1994, 1995; Gass, 1997; Gass & Varonis, 1994). Some of the processes involved in output production would be repetition and form adjustments, which facilitates the noticing function (Færch & Kasper, 1987; Long, 1996). According to Pica (1994), this form of interaction and output production may be regarded as an attention focusing mechanism (de la Fuente, 2002) if language learners do take heed of the feedback given by the instructors. By interpreting this in the light of the aforementioned cognitive theories, this may probably shed light on why the participants from the GPIO group, who inferred meaning from the available contextual clue given without needing to pay attention to the form of new word, did not perform as well as the GINP group (Coady, 1993; R. Ellis, 1994).

The second peripheral hypothesis predicted that there is not significant difference between learners who are exposed to negotiated input with output production than those who are exposed to negotiated input without output in relation to receptive vocabulary acquisition is not confirmed by the study. Encouraging results revealed that there was a significant difference between the GINP and GINW group in relation to receptive vocabulary acquisition. The findings of the study corroborate with the study conducted by Ellis and He (1999), whereby the output group outscored the negotiated input group. On the other hand, the results contrasted the findings published by Loschky (1994) and de la Fuente (2002), who revealed no significant difference between the output group and the interactionally modified group. This result may be explained through Swain's (1985) study, in which participants who attended language immersion

classes faced communication barrier despite being placed in an input-rich environment. This was primarily due to the lack of opportunity for comprehensible output. Therefore, Swain (1985, 1995) suggested that learners should be provided with negative feedback in order to force learners to make choices in producing meaningful message, which are similar to the conditions provided in the GINP group. Akin to Lantolf and Pavlenko's (1995) contention, Swain (1985) argues that second language acquisition does not solely rely on the richness of input, but rather to what extent the learner takes control over his or her own learning by paying attention to new words. When learners are engaged in producing output, they are involved in deeper mental processes of semantic and syntactic processing (Swain, 1895). Furthermore, Sharwood Smith (1987) pointed out that receptive acquisition gradually progresses towards productive mastery, which may explain the need of output production in L2 receptive vocabulary acquisition. Therefore, it can be observed that based on this study, the absence of output production in the GINW group yielded lower level of receptive vocabulary acquisition.

5.2.3 Hypothesis 3

The third hypothesis predicted that learners who are exposed to pre-modified input with negotiated interaction with or without output production will attain higher level of productive vocabulary acquisition than those learners who are exposed to pre-modified input only is partially confirmed in this study. The results of the study revealed that there was no significant difference between the GPIO and GINW group. However, the findings revealed that there was a significant difference between the GPIO and GINP group, in which the GINP group managed to outperform the GPIO group. The current study supported previous studies by Ellis et. al (1994), Ellis and He (1999) and de la Fuente (2002), in which suggested that the presence of output production accompanied by negotiation would facilitate L2 productive vocabulary acquisition. Ellis et. al (1994) asserted that when learners are provided with negative

feedback from the instructors and are pushed to produce in the L2, only then negotiation would be able to promote productive acquisition of new words. These empirical studies are grounded under the basis of Swain's output hypothesis (1985) that producing output is pertinent where new lexical items are concerned. The results of the present study supported the results obtained by Ellis and He (1999), who found that the output production group produced significantly more words than the pre-modified input group. Their study also revealed that there was no significant difference between the negotiated input group and the pre-modified input group. Additionally, the results may be able to confirm studies conducted on looks specifically into the role of output in L2 vocabulary acquisition (Ellis et al. 1994; Gass & Varonis, 1994; Mackey & Philp, 1998; Pica et al., 1989) or directly (Mackey, 1995, 1997, 1999; Swain & Lapkin, 1995; Van de Branden, 1997).

In response to Krashen's (1985) comprehensible input, Swain (1985, 1995) postulated that comprehensible input may be beneficial for L2 comprehension but it remains insufficient for L2 acquisition (Swain & Lapkin, 1995). The empirical findings of this study seem to substantiate that input and output promotes L2 acquisition through different processes with respect to L2 vocabulary acquisition (Swain & Lapkin, 1995). As discussed earlier on for the second hypothesis, output production encourages language learners to perform specific tasks by paying attention and notice the gap, which required for production of new words in the L2.

The third peripheral hypothesis states that there is no significant difference between learners who are exposed to negotiated input without output production and learners who are exposed to negotiated input with output production in relation to productive vocabulary acquisition is not confirmed in this study. It appears that there was a statistical difference between the Test 2 and Test 3 of the GINP group, whereby the participants produced more words significantly in the final test. The results

confirmed the studies conducted by Ellis and He (1999) and Gass and Varonis (1994). Based on the empirical findings of the studies mentioned, it was observed that production in the L2 occurred at a delayed time during communicative events, which suggested that output production was beneficial for delayed retention but not immediate retention. This further substantiates the findings by Gass (1997) that puts forth the notion that the impact of negotiation has significant effects on language development among language learners in a delayed time frame.

The results also revealed that the groups under three types of negotiation attained reasonable level of productive vocabulary acquisition and retention. However, the GINP group experiences a significant increase in test scores from Test 2 to Test 3 in relation to PVKS. It may be argued that this may be due to test effect and input exposure between tests, where the participants increase their exposure to the target items from one test to another. However, it is interesting to note that there was still a significance difference between the test scores obtained by the GINW and GINP group as all three groups had equal opportunity to test and input exposure. According to Gass and Varonis (1994), the significance drawn from the results may be attributed by time factor. Gass and Varonis (1994) further argued that although interlanguage development may not occur within a short-term, it might be possible for long-term retrieval because learners are given the time to internalize and utilized the feedback obtained. Similarly, Mackey (1999) demonstrated that retention in the long-term is applicable for other linguistic element such as grammatical development. Mackey (1999) asserted that time is required to internally process and integrate the treatment to their current linguistic knowledge. The contention of short-term and long-term effects of L2 vocabulary acquisition or any other linguistic areas may be scrutinized for future researches. In the light of cognitive theories, it appears that the findings of this present study on the statistical differences between receptive and productive vocabulary acquisitions test

scores supports the above notion. Both receptive and productive acquisition entails different mental processes; receptive acquisition requires the learners to detect and identify the words where as productive acquisition denotes deeper input processes of internalizing and integrating the words as part of their vocabulary. Nonetheless, when the participants are pushed to produce in the L2, it can be seen that encouraging results emerged and is proven beneficial for the process of internalization and utilization of the target words in the long term.

The notion of attention (Swain, 1985) enriched by Long (1990) and Schmidt, (1990, 1993) has been closely associated to the comprehensible input and comprehensible output in the field of SLA. Data from several sources have revealed the pertinence of attention in language development (Leow, 1997, 1998, 2000; Rosa & O'Neill, 1999). These studies have attempted to confirm that attention ensures learners to pay notice at their linguistic gaps and eventually transform input into intake. Based on the interactions drawn from the observation, input is made comprehensible with the negotiated interaction that took place between the instructor and the participants in GINW group.

(2) Interaction (Task 1)

Instructor: Find the carrot.
Participant: Is purple in colour?
Instructor: No, it's not purple.
Participant: Is it round?
Instructor: No, it's not round.
Participant: Oh, it is long and thin yah.
Instructor: Yes, you are right.

(3) Interaction (Task 2)

Instructor: Put the onion below the table.
Participant: Onion? What smell?
Instructor: A strong smell. It will make you cry.
Participant: Cry?
Instructor: Yes, it stings your eyes when you cut it.
Participant: Can eat just like that?
Instructor: Yes, sometimes you can eat it just like that with satay.
Participant: Oh, okay.

As portrayed in interaction (2) and (3), the participants underwent conscious effort in noticing the gap in their current linguistic knowledge. In the process, when the participants pose series of questions, attention is directed to the lexical item. The instructor assumed a salient role in facilitating the noticing the gap process by providing the required information to the participants in understanding the target lexical items. This will prompt the participants to comprehend and decode the input received. It is important to note that the participants from the GINW group did not repeat or understand how the target words function in sentence structures. In line with Long's (1990, 1995) interaction hypothesis, input is made comprehensible through the act of negotiation. As claimed by Sharwood Smith (1987), receptive acquisition gradually progresses towards productive acquisition. Comprehension of input remains as the first stage for the language learners to recognize the word and to propel learners to notice how words function in forms, production of new linguistic forms may be required to internalize, integrate and activate the use of the word knowledge. As a result, it can be observed that GINP group produces and interacts in forms rather than the GINW group.

Swain (1985) clearly outlines the two functions linked to output production in SLA, which are (a) noticing the gap and (b) hypothesis testing. Relevant to this study, pre-modified input that has been made comprehensible decreases consciousness-raising among language learners, which may not stimulate learners to notice their linguistic gaps. However, production of output may compensate this drawback. When learners are provided the room to formulate and test out different expressions to communicate, this increases awareness and ensures that learners pay attention to the existing gap in L2. Several studies have examined the importance of attention with regards to L2 acquisition, in which demonstrated positive correlation between levels of awareness and levels of language acquisition (Leow, 1997; Rosa & O'Neill, 1999). The empirical studies cited essentially incorporated hypothesis testing and multiple mental processes

to increase the level of awareness. This will explicitly inform language teachers that in order for young learners to acquire a second language, they should be forced to produce any linguistic forms. This view is supported by the fact that the GINP group outperformed the groups with other conditions in relation to receptive and productive acquisition. Participants from the GINP group are pushed to produce instructions in the L2 and in the process; they would need to repeat the words. When learners are forced to engage in such circumstances, they may indirectly develop their interlanguage by integrating new lexical items to their current linguistic knowledge (Larsmen-Freeman & Long, 1991). The above notion can be demonstrated through the examples from the interactions observed in the GINP group for the second round of each task.

(4) Interaction (Task 1)

Participant: Hmm.. What is it? Look like green tree. I eat it with carrot.

Instructor: Do you remember the name of the vegetable?

Participant: Mmm.. Take the broco..

Instructor: Almost right. Broccoli.

Participant: Yah. Broccoli.

(5) Interaction (Task 2)

Participant: Put the thing.. that is green..

Instructor: Can you remember the name of the vegetable?

Participant: Hmm.. broccoli

Instructor: Broccoli? This one? Where do I put the broccoli?

Participant: Yes. That one! Put the broccoli on the chair.

Interaction (4) witnessed that the participant noticed the gap in his linguistic knowledge as he struggled to recall the lexical item tested by different means. As argued by Gass and Varonis (1994), one of the methods used by young learners to recall the lexical item needed for output production is through physical referral (“look like green tree.”) or meanings of the word (“I eat it with carrot.”). After being prompted by the instructor, the participant underwent the hypothesis testing stage by attempting to produce the lexical form (“Take the broco..”) albeit incomplete. This stage is argued to be critical as it functions akin to a catalyst as it thrusts the participant to acquire the

intended lexical form. Interaction (5) records and shows the progress made by the same participant. Similar to interaction (4), it can be observed that the participant switches to indirect physical referral of the lexical item. The pertinent role played by the instructor was a determinant factor to force the participant to produce the lexical form in L2. The participant managed to produce the actual lexical form without any pronunciation mistake and with some assistance and feedback from the instructor; the participants produced the lexical item in a sentence structure. Although the participant was not able to produce the lexical item in the immediate PVKS, he managed to produce the word in both delayed PVKS instead. In line with Swain and Lapkin's (1995) view, output production in the GINP may have facilitated L2 vocabulary acquisition from imperfect generation of lexical items ("brocco") to usage of word in actual forms ("Put the broccoli on the chair.").

5.3 Summary and Implications of the Study

From the hypotheses formulated in the study, it can be concluded that:

1. Learners who are given the opportunity to negotiate for meaning attained higher level of vocabulary comprehension than those learners who are exposed to pre-modified input only.
2. Learners who are given the opportunity to negotiate and produce output attained higher level of receptive vocabulary acquisition than those learners who are exposed to pre-modified input only. Apart from that, there were significant effect between learners who are exposed to negotiation and output and learners who are given the chance to negotiate only.
3. Learners who are given the opportunity to negotiate and produce output attained higher level of productive vocabulary acquisition than those learners who are exposed to pre-modified input only. Similar to receptive acquisition, output

production appears to have differential effects on productive vocabulary acquisition than learners who are given the chance to negotiate only.

Findings from the present study confirmed the positive effects of the existing interaction hypothesis framework specifically in facilitating the role of comprehensible output (Swain, 1985, 1995) and comprehensible input (Krashen, 1985) in L2 vocabulary comprehension, receptive and productive vocabulary acquisition. Enriched with the works of Long (1996) and Gass (1997), hypothesis testing and noticing the gap as a result of attention may be one of the determinant factors in the interaction to promote L2 vocabulary development.

The data findings from the three research questions explicitly explore the effects of three different input conditions in L2 vocabulary acquisition among primary ESL learners in Malaysia. Based on this experimental study, negotiation and output production have been proven to be beneficial for the experimental groups to acquire lexical items than those in the control group. However, having gone through 11 years of formal lessons in English, one may wonder what are some of the reasons why Malaysian graduates are unable to attain reasonable level of communication skills in English. This may be attributed to current exam-oriented classroom culture in Malaysia. Language teachers, especially those who are serving in primary national schools, are often thrown with an ultimatum; in order to yield more A's in the pupils' examination slips, knowledge-based examinations and form-based tasks are given priority. This in turn may neglect listening and speaking skills, which are the skills required for basic communication. Therefore, the evidence from this study suggests that teachers should provide opportunities for pupils to interact and practice the target language in ESL classrooms based on the valuable benefits offered by interactive tasks. Interactive activities have also been regarded as unfeasible due to time constraints as language teachers are pressured to complete the syllabus in time. Thus, the results of this research

support the idea that school authorities should acknowledge the pertinent role of negotiation in the field of SLA, which would alleviate the mentioned pressure. Aside from gearing the pupils up for public examinations, they are presented to opportunity to experience interesting and enjoyable lessons when language teachers are given the authority to conduct interactive activities as part of the syllabus. Through negotiated interaction, fun and interactive tasks in ESL classrooms will help to ensure that motivation is sustained to boost participation among young learners to practice in a language that is deemed intimidating and foreign for language development. Having said that, instead of an ultimatum, equal importance should be given to form-based tasks, knowledge-based examinations and interactive tasks. In line with the National Education Philosophy, learners gain greater advantage in the competitive world of science, technology, economics, and professional sectors by equipping themselves not only with written knowledge but also, communication skills in English.

5.4 Recommendations for Future Research

However, a number of caveats need to be noted regarding the present study. Firstly, the current research was not specifically designed to evaluate variables related to other word classes and more complex aspect of lexical acquisition. The target items tested focused on a specific word class, a list consisting of 10 concrete nouns in the English Language. The selection of the nouns, which comprises of different types of vegetables, is based on the topical word lists to be learnt as stipulated in the Standard Document (Curriculum Development Centre, 2011). The study will investigate the acquisition of the base forms and meaning of the nouns. Therefore, further research should be done to investigate the other word classes (eg. verbs, adjectives, adverbs) and other aspects of lexical acquisition (eg. associations, collocations).

Apart from that, the present study utilizes one-way interactive tasks to examine L2 vocabulary comprehension and acquisition that yielded positive results. In both the

listening comprehension tasks, the information exchanged was one way as only the teacher holds the information. Previous works on negotiated interaction in SLA have suggested that two-way interactive tasks may be beneficial for classroom settings by providing ample room for interactional modifications. Thus, a future study employing information gap tasks between pupils in pairs would be very interesting. In addition to that, the artificial nature of the samples selected comprised of pupils who are of similar level of proficiency in English. This may raise threats against ecological validity when issues pertaining to shifting theories to real classrooms are addressed. Therefore, it is recommended that further research be undertaken to include pupils who are of different levels of proficiency such as those who are of average and intermediate levels.

Thirdly, the current study was unable to investigate the cognitive processes that occurred after feedback during negotiated interactions (Gass, 1997; Long, 1996; Mackey, 1997) and how that spurred L2 vocabulary development. Hence, the grey area of feedback in negotiation is an intriguing one, which could be usefully explored in further research. As recommended in earlier researches, a mixed methodology that fuses quantitative and qualitative data analyses may aid in understanding grammar development (Mackey, 1997) and vocabulary development (Gass, 1997; Long, 1996). Likewise, future research is also needed to determine the cognitive processes involved in output production on L2 vocabulary development. Although the findings of the study demonstrated positive effects of modified output, how language learners are propelled to produce the output and the processes that are responsible to transform input into output remains unknown. Further investigation and experimentation into output as the sole independent variable is strongly recommended (de la Fuente, 2002)

It is also important to note that with a small sample size, caution must be applied, as the findings might not be transferable to other language learners of different educational settings. Larger samples of participants could provide more definitive

evidence. A number of possible future studies using the same experimental set up are apparent. With that notion in mind, multiple assessments may be adapted to measure vocabulary comprehension and acquisition to enhance validity and reliability of the study.

5.5 Conclusion

This chapter has presented an overview of the significant impact of negotiated interaction among primary ESL learners by engaging the role of input and output for L2 vocabulary acquisition. The data analysis and discussion clearly displayed the positive outcomes of negotiated interaction in an ESL classroom. In relation to theoretical implications, the study may contribute to another literature of negotiated interaction and output production. The findings of the study may support its effectiveness towards L2 vocabulary acquisition, in association to nouns, among primary ESL learners in Malaysia. More importantly, in order to address the current problem of increasing number of unemployed graduates in Malaysia, communicative activities may be regarded as a commodity to improve the speaking skills of young ESL learners.

List of References

- Abdul Fatah Abdul Malik & Mohd. Majid Konting. (1993). *Memahami penyelidikan pendidikan : Satu pengenalan* (Edisi Keempat). Selangor, Malaysia : Universiti Putra Malaysia.
- Aliaga, M., & Gunderson, B. (2000). *Interactive statistics*. Saddle River, NJ: Prentice Hall.
- Anastasi, A.(1982). *Psychological testing* (5th ed.). New York, NY: Macmillan.
- Arnheim, Rudolf. (1974). *Art and visual perception*. Berkeley, CA: University of California Press.
- Blachowicz, C.L.Z., & Fisher, P.J. (2006). *Teaching vocabulary in all classrooms* (3rd ed.). Upper Saddle River, NJ: Pearson Education.
- Beck, I., & McKeown, M. (1991). Conditions of vocabulary acquisition. In R. Barr, M. Kamil, P. Mosenthal & P. D. Pearson (Eds.), *Handbook of reading research* (Vol. 2, pp. 789-814). New York, NY: Longman.
- Bitchener, J. (2003). Does corrective feedback help L2 learners improve the accuracy of their writing over time? *New Zealand Studies in Applied Linguistics*, 9 (2), 73-87.
- Bland, J. M., & Altman, D. G. (1990). A note on the use of the intra-class correlation coefficient in the evaluation of agreement between two methods of measurement. *Computers in Biology and Medicine*, 20, 337–340.
- Breen, M. P., & Littlejohn, A. (2000). *Classroom decision making*. Cambridge, England: Cambridge University Press.
- Brown, C. (1993). Factors affecting the acquisition of vocabulary: Frequency and saliency of words. In T. Huckin, M. Haynes & J. Coady (Eds.), *Second language reading and vocabulary learning* (pp. 263–286). Norwood, NJ: Ablex.

- Bryman, A. (2012). *Social research methods* (4th ed.). Oxford, England: Oxford University Press.
- Carpenter, S. K., & Olson, K. M. (2012). Are pictures good for learning new vocabulary in a foreign language? Only if you think they are not. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38, 92–101.
- Carter, R. (1987). *Vocabulary: Applied linguistics perspectives*. London: Routledge.
- Chapelle, C.A. (1998). Construct definition and validity inquiry in SLA research. In L. F. Bachman & A. D. Cohen (Eds.), *Interfaces between second language acquisition and language testing research* (pp. 32–70). Cambridge, England: Cambridge University Press.
- Charles, C. M., & Mertler, C. A. (2002). *Introduction to educational research* (4th ed.). Boston, MA: Allyn & Bacon.
- Chaudron, C. (1983). Simplification of input: Topic reinstatements and their effects on L2 learners' recognition and recall. *TESOL Quarterly*, 17 (3), 437-458.
- Chomsky, N. (1965). *Aspects of the theory syntax*. Cambridge, MA: MIT Press.
- Coady, J. (1993). Research on ESL/EFL vocabulary acquisition: Putting it in context. In T. Huckin, M. Haynes, & J. Coady (Eds.), *Second language reading and vocabulary learning* (pp. 16–28). Norwood, NJ: Ablex.
- Cohen, A.D. (1990). *Language learning: Insights for learners, teachers, and researchers*. Boston, MA: Heinle & Heinle.
- Cook, V. (1991). Universal Grammar and second language acquisition. *System*, 19, 342–347.
- Corder, S.P. (1967). The significance of learners' errors. *International Review of Applied Linguistics*, 5, 161-169
- Cozby, P.C. (2001). *Measurement concepts: Methods in behavioral research* (7th ed.). California, CA: Mayfield Publishing Company.

- Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning & Verbal Behavior*, 11, 671-684.
- Creswell, J. W. (2005). *Educational research: Planning, conducting and evaluating quantitative and qualitative research* (2nd ed.). Upper Saddle River, NJ: Pearson Education.
- Curriculum Development Centre. Ministry of Education, (2012). Performance *standard document KSSR English language year 3*. Retrieved from KSSR Online website: <https://onedrive.live.com/view.aspx?resid=E282FD7928D2FF3!706&app=Word>
- Curriculum Development Centre. Ministry of Education, (2011). *Standard document KSSR English language*. Retrieved from KSSR Online website: <https://skydrive.live.com/?cid=0e282fd7928d2ff3&id=E282FD7928D2FF3!605>
- Dale, E. (1965). Vocabulary measurement: Techniques and major findings. *Elementary English*, 42(8), 895–901.
- Davidshofer, C. O., & Murphy, K. R. (2005). *Psychological testing: Principles and applications* (6th ed.). Upper Saddle River, N.J.: Pearson / Prentice Hall.
- de la Fuente, M. J. (2002). Negotiation and oral acquisition of L2 vocabulary: The roles of input and output in the receptive and productive acquisition of words. *Studies in Second Language Acquisition*, 24, 81-112.
- de la Fuente, M. J. (2006). Classroom L2 vocabulary acquisition: investigating the role of pedagogical tasks and form-focused instruction. *Language Teaching Research*, 10 (3), 263–295.
- Denzin, N. K. (1989). *The research act* (3rd ed.). Englewood Cliffs, NJ: Prentice Hall.
- Doane, D.P., & Seward, L. E. (2011). Measuring skewness: A forgotten statistic? *Journal of Statistics Education*, 19 (2).
- Doughty, C., & Pica, T. (1986). 'Information-gap' tasks: Do they facilitate second language acquisition? *TESOL Quarterly*, 20(2), 305-326.

- Cramer, D., & Howitt, D. (Eds.). (2004). *The SAGE dictionary of statistics*. London, England: SAGE Publications, Ltd.
- Elley, W. B. (1989). Vocabulary acquisition from listening to stories. *Reading Research Quarterly*, 24, 174-186.
- Ellis, N., & Beaton, A. (1993). Psycholinguistics determinants of foreign language vocabulary learning. *Language Learning*, 43, 559-617.
- Ellis, R. (1991). The interaction hypothesis: A critical evaluation. In E. Sadtono (Ed.), *Language acquisition and the second/foreign language classroom*. Singapore: RELC.
- Ellis, R. (1994). Factors in the incidental acquisition of second language vocabulary from oral input: A review of essay. *Applied Language Learning*, 5, 1-32.
- Ellis, R. (1995). Modified oral input and the acquisition of word meanings. *Applied Linguistics*, 16, 409-435.
- Ellis, R. & Heimbach, R. (1997). Bugs and birds: Children's acquisition of second language vocabulary through interaction. *System*, 25, 247-259.
- Ellis, R., Heimbach, R., Tanaka, Y., & Yamazaki, A. (1999). Modified input and the acquisition of word meanings by children and adults. In R. Ellis (Ed.), *Learning a second language through interaction* (pp. 63-114). Amsterdam/Philadelphia, PA: John Benjamins Publishing Company.
- Ellis, R., & He, X. (1999). The roles of modified input and output in the incidental acquisition of word meanings. *Studies in Second Language Acquisition*, 21, 285-301.
- Ellis, R., Tanaka, Y., & Yamakazi, A. (1994). Classroom interaction, comprehension, and L2 vocabulary acquisition. *Language Learning*, 44, 449-491.
- Erlam, R. & Loewen, S. (2010). Implicit and explicit recasts in L2 oral French interaction. *The Canadian Modern Language Review*, 67(4), 877-905.

- Faerch, C., Haastруп, K., & Phillipson, R. (1984). *Learner language and language learning*. Clevedon, England: Multilingual Matters.
- Federer, W.T. (1955). *Experimental design - Theory and application*. New York, NY: MacMillan Publishing Co. Inc.
- Fisher, D., Brozo, W.G., Frey, N., & Ivey, G. (2011). *50 instructional routines to develop content literacy* (2nd ed.). Boston, MA: Pearson.
- Foster, P. (1998). A classroom perspective on the negotiation of meaning. *Applied Linguistics*, 19, 1–23.
- Fraenkel, J. R., and Wallen, E. N. (1996). *How to design and evaluate research*. Ohio, OH : McGraw-Hill, Inc.
- Gall, M. D., Borg, W. R., & Gall, J. P. (2003). *Educational research: An introduction* (7th ed.). White Plains, NY: Longman.
- Garfinkel, H. (1967). *Studies in ethnomethodology*. Eaglewood Cliffs, NJ: Prentice-Hall.
- Gass, S. (1997). *Input, interaction, and the second language learner*. Mahwah, NJ: Erlbaum.
- Gass, S., & Madden, C. (Eds). (1985). *Input in second language acquisition*. Rowley, MA: Newbury House.
- Gass, S., & Torres, A.M. (2005). Task-based interactions in classroom and laboratory settings. *Language Learning (The best of Language Learning Series)*, 61, 189-220.
- Gass, S., & Varonis, M. (1994). Input, interaction, and second language production. *Studies in Second Language Acquisition*, 16, 283–302.
- George, D., & Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference. 11.0 update* (4th ed.). Boston, MA: Allyn & Bacon.

- Gibson, J. J. (1969). *Principles of perceptual learning and development*. New York, NY: Appleton-Century-Crofts.
- Gibson, J. J. (1971). The information available in pictures. *Leonardo*, 4, 27-35.
- Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston, MA: Houghton Mifflin Company.
- Gregg, K.R. (1984). Krashen's Monitor and Occam's razor. *Applied Linguistics*, 5, 79-100.
- Guba E.G., & Lincoln Y.S. (1981). *Effective evaluation: Improving the usefulness of evaluation results through responsive and naturalistic approaches*. San Francisco, CA: Jossey-Bass.
- Gumperz, J. (1971). *Language in social groups*. Stanford, CA: Stanford University Press.
- Haladyna, T. M. (1999). *Developing and validating multiple choice test items* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
- Hall, R. (1998). *Extraneous and Confounding Variables and Systematic vs non-systematic error*. Retrieved from <http://web.mst.edu/~psyworld/extraneous.htm>
- Hartley, J., and Chesworth, K. (2000). Qualitative and quantitative methods in research on essay writing: No one way. *Journal of Further and Higher Education*, 24 (1), 15-24.
- Hatch, E. (Ed). (1978). *Second language acquisition: A book of readings*. Rowley, MA: Newbury House.
- Hatch, E. (1983). Simplified input and second language acquisition. In R. Andersen (Ed.), *Pidginization and creolization in second language acquisition* (pp. 64-86). Rowley, MA: Newbury House.

- Hatch, E., & Brown, C. (1995). *Vocabulary: Semantics and language education*. Cambridge, England: Cambridge University Press.
- Harmer, J. (2001). *The practice of English language teaching*. Essex, England: Longman Press.
- Herron, C., Hanley, J., & Cole, S. (1995). A comparison study of two advance organizers for introducing beginning foreign language students to video. *Modern Language Journal*, 79 (3), 387-394.
- Hockley, W. E. (1994). Reflections of the mirror effect for item and associative recognition. *Memory & Cognition*, 22, 713-722.
- Huck, S. W., & McLean, R. A. (1975). Using a repeated measures ANOVA to analyze data from a pretest-posttest design: A potentially confusing task. *Psychological Bulletin*, 82, 511- 518.
- Hudson, T. (1982). The effects of induced schemata on the 'short circuit' in L2 reading: Non-decoding factors in L2 reading performance. *Language Learning*, 32, 1-31.
- Hughes, P. (2001). Paradigms, methods and knowledge. In G. MacNaughton, S. R. Rolfe & I. Siraj-Blatchford (Eds.), *Doing early childhood research: International perspectives on theory and practice* (pp. 31-55). NSW, Australia: Allen & Unwin.
- Hulstijn, J. H., & Hulstijn, W. (1984). Grammatical errors as a function of processing constraints and explicit knowledge. *Language Learning*, 34, 23-43.
- Joe, A. (1994). *Generative use and vocabulary learning*. (Unpublished master's thesis). Victoria University of Wellington, NZ.
- Johnson, R.B., & Christensen, L.B. (2000). *Educational research: Quantitative and qualitative approaches*. Boston, MA: Allyn and Bacon.
- Keppel, G. (1991), *Design and analysis: A researcher's handbook* (3rd ed.). Englewood Cliffs, NY: Prentice Hall.

- Keppel, G., & Wickens, T. D. (2004). *Design and analysis: A researcher's handbook* (4th ed.). Upper Saddle River, NJ: Pearson.
- Kirk, R. E. (1982). *Experimental design: Procedures for the behavioral sciences*. Monterey, CA: Brooks/Cole.
- Kramsch, C. (1995). Rhetorical models of understanding. In T. Miller (Ed.), *Functional approaches to written texts: Classroom applications* (pp 61-78). Paris: USIS.
- Krashen, S. (1980). *Second language acquisition and second language learning*. Oxford, England: Pergamon Press.
- Krashen, S. (1982). *Principles and practice in second language acquisition*. Oxford, England: Pergamon Press.
- Krashen, S. (1985). *The input hypothesis: Issues and implications*. London, England: Longman.
- Kuder, G. F., & Richardson, M. W. (1937). The theory of the estimation of test reliability. *Psychometrika*, 2, 151–160.
- Kütük, R. (2007): The effect of mnemonic vocabulary learning strategy and storytelling on young learners' vocabulary learning and retention. (Unpublished master's thesis). Cukurova University, Turkey.
- Landis, J.R., & Koch, G.G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33, 159–174.
- Landsberger, H. A. (1958). *Hawthorne Revisited*. Ithaca, NY: Cornell University.
- Lantolf, J. P., & Pavlenko, A. (1995). Sociocultural theory and second language acquisition. *Annual Review of Applied Linguistics*, 15, 108-124.
- Larsen-Freeman, D., & Long, M. H. (1991). *An introduction to second language acquisition research*. New York, NY: Longman.

- Laufer, B. (1990). Words you know: How they affect the words you learn. In J. Fisiak (Ed.), *Further insights into contrastive linguistics* (pp. 573–593). Amsterdam, Netherlands: John Benjamins.
- Laufer, B. (1998). The development of passive and active vocabulary in a second language: Same or different? *Applied Linguistics*, 19, 255–271.
- Laufer, B., & Goldstein, Z. (2004). Testing vocabulary knowledge: size, strength and computer adaptiveness. *Language Learning* 54 , 399-436.
- Lee, J. 2000. *Tasks and communicating in language classrooms*. New York, NY: McGraw-Hill.
- Leow, R. P. (1997). Attention, awareness, and foreign language behavior. *Language Learning*, 47, 467–506.
- Leow, R. P. (1998). Toward operationalizing the process of attention: Evidence for Tomlin and Villa's (1994) fine-grained analysis of attention. *Applied Psycholinguistics*, 19, 133–159.
- Leow, R. P. (2000). A study of the role of awareness in foreign language behavior: Aware versus unaware learners. *Studies in Second Language Acquisition*, 22, 557–584.
- Levin, R. I., & Rubin, D.S. (2000). *Statistics for management* (7th ed.). Englewood Cliffs, N.J: Prentice-Hall.
- Lewis, M. (1997). *Implementing the lexical approach: Putting theory into practice*. Hove, England: Language Teaching Publications.
- Li, K., & Hudak, P. (1989). Memory coherence in shared virtual memory systems. *Transactions on Computer System (ACM)*, 7 (4), 321-59.
- Long, M. H. (1981). Input, interaction, and second language acquisition. In H. Winitz (Ed.), *Native language and foreign language acquisition: Annals of the New York academy of sciences* (pp. 259–278). New York: NY: Academy of Sciences.

- Long, M. H. (1983). Native speaker/non-native speaker conversation in the second language classroom. In M. Clark, M., & Handscombe, J. (eds.), *On TESOL '82: Pacific perspectives on language learning* (pp. 207-25). Washington, D.C., WA: TESOL.
- Long, M. H. (1985). Input and second language acquisition theory. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 377–393). Rowley, MA: Newbury House.
- Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413–468). San Diego, CA: Academic Press.
- Loschky, L. (1994). Comprehensible input and second language acquisition: What is the relationship? *Studies in Second Language Acquisition*, 16, 303–324.
- Lyster, R., & Izquierdo, J. (2009). Prompts versus recasts in dyadic interaction. *Language Learning*, 59(3), 458.
- Mackey, A., Adams, R., Stafford, C., & Winke, P. (2010). Exploring the relationship between modified output and working memory capacity. *Language Learning*, 60, 501-533.
- Mackey, A. (1997). *Stepping up the pace: An empirical study of input, interaction, and second language development*. (Unpublished manuscript). Michigan State University, East Lansing.
- Mackey, A., Gass, S., & McDonough, K. (2000). How do learners perceive interactional feedback? *Studies in Second Language Acquisition*, 22, 471-497.
- Mackey, A., & Philp, J. (1998). Conversational interaction and second language development: Recasts, responses, and red herrings? *The Modern Language Journal*, 82, 338-356.

- Maiguashca, R. U. (1993). Teaching and learning vocabulary in a second language: Past, present and future directions. *The Canadian Modern Language Review*, 50 (1), 83-100.
- Martin, W. E., & Bridgmon, K.D. (2012). *Quantitative and statistical research methods: From hypothesis to results*. California, CA: Jossey-Bass.
- Maxwell, S. E., Delaney, H. D., & Manheimer, J. M. (1985). ANOVA of residuals and ANCOVA: Correcting an illusion by using model comparisons and graphs. *J. Educ. Statist.*, 10, 197-209.
- McMillan, J. H., & Schumacher, S. (2001). *Research in education: A conceptual introduction* (5th ed.). New York, NY: Longman.
- McNeill, M. (1996). Vocabulary knowledge profiles: Evidence from Chinese-speaking ESL teachers. *Hong Kong Journal of Applied Linguistics*, 1, 39-63.
- Meara, P. (1980). Vocabulary acquisition: A neglected aspect of language learning. *Language Teaching and Linguistics*, 15 (4), 221 – 46.
- Meara, P. (1996). The dimensions of lexical competence. In G. Brown, K. Malmkjaer, & J. Williams (Eds.), *Performance and competence in second language acquisition* (pp. 35–53). New York, NY: Cambridge University Press.
- Melka Teichroew, F.J. (1982). Receptive vs. productive vocabulary: a survey. *Interlanguage Studies Bulletin*, 6, 5-33.
- Mills, G. E. (2003). *Action research: A guide for the teacher researcher* (2nd ed.). Upper saddle River, NJ: Merrill/Prentice-Hall.
- Milton, J. (2009). *Measuring second language vocabulary acquisition*. Clevedon, England: Multilingual Matters.
- Mueller, G. (1980). Visual contextual clues and listening comprehension: An experiment. *Modern Language Journal*, 64, 335-40

- Nabei, T. (2002). *Recasts in classroom interaction: A teacher's intention, learners' awareness, and second language learning*. (Unpublished doctoral dissertation). University of Toronto (OISE), Canada.
- Nation, P. (1982). Beginning to learn foreign vocabulary: A review of the research. *RELC Journal*, 13, 14-36.
- Nation, P. (1990). *Teaching and learning vocabulary*. Rowley, MA: Newbury House.
- Nation, P. (2001). *Learning vocabulary in another language*. Cambridge, England: Cambridge University Press.
- Nelson, D. L., Reed, V. S., & McEvoy, C. L. (1977). Learning to order pictures and words: A model of sensory and semantic encoding. *Journal of Experimental Psychology: Human Learning & Memory*, 3, 485-497.
- Nelson, D. L., Reed, V. S., & Walling, J. R. (1976). Pictorial superiority effect. *Journal of Experimental Psychology: Human Learning & Memory*, 2, 523-528.
- Ng, E. (2013, May 27). English emphasis lacking in public schools. *The Star*. Retrieved from <http://www.thestar.com.my/News/Nation/2013/05/27/English-emphasis-lacking-in-public-schools.aspx>
- Ng, L.L., & Sheila Marina Sappathy. (2011). L2 vocabulary acquisition: The impact of negotiated interaction. *GEMA Online™ Journal of Language Studies*, 11(2). Retrieved from [http://www.ukm.my/ppbl/GEMA%20vol%2011%20\(2\)%202011/abstract%205-20.pdf](http://www.ukm.my/ppbl/GEMA%20vol%2011%20(2)%202011/abstract%205-20.pdf)
- Nor Hashimah Jalaludin, Norsimah Mat Awal & Kesumawati Abu Bakar. (2008). The mastery of English language among lower secondary school students in Malaysia: A linguistic analysis. *European Journal of Social Sciences*, 7 (2), 106-119.

- Nornadiah Razali, & Wah, Y. B. (2011). Power comparison of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson Darling tests. *Journal of Statistical Modeling and Analytics*, 2 (1), 21-33.
- Nunan, D. (1989). *Designing tasks for the communicative classroom*. Cambridge, England: Cambridge University Press.
- Nunan, D. (1991). *Language teaching methodology*. London, England: Prentice Hall International.
- O'Grady, W. (1998). The syntax of idioms. *Natural Language and Linguistic Theory*, 16, 79-312.
- Oliver, R. (1998). Negotiation of meaning in child interaction. *The Modern Language Journal*, 82, 372-386.
- Omaggio, A. (1979). Pictures and second language comprehension: Do they help? *Foreign Language Annals*, 12, 107-16.
- Pak, J. (2012, March 5). *Why some parents are worried about the decline of english in malaysia*. Retrieved from <http://pri.org/stories/2012-03-05/why-some-parents-are-worried-about-decline-english-malaysia>
- Palmberg, R. (1987). Patterns of vocabulary development in foreign language learners. *Studies in Second Language Acquisition*, 9, 201-220.
- Paivio, A., & Csapo, K. (1973). Picture superiority in free recall: Imagery or dual coding? *Cognitive Psychology*, 5, 176-206.
- Paivio, A., Rogers, T. B., & Smythe, P. C. (1968). Why are pictures easier to recall than words? *Psychonomic Science*, 11, 137-138.
- Pica, T. (1992). The textual outcomes of native speaker–non-native speaker negotiation: What do they reveal about second language learning? In C. Kramsch & S. McConnell-Ginet (Eds.), *Text and context: Cross-disciplinary perspectives on language study* (pp. 198–237). Lexington, MA: Heath.

- Pica, T. (1994). Research on negotiation: What does it reveal about second language learning conditions, processes, and outcomes? *Language Learning*, 44, 493–527.
- Pica, T., Young, R., & Doughty, C. (1987). The impact of interaction in comprehension. *TESOL Quarterly*, 21, 737–758.
- Pica, T., Kanagy, R., & Falodun, J. (1993). Choosing and using communication tasks for second language instruction. In G. Crookes & S. Gass (Eds.), *Tasks and language learning: Integrating theory and practice* (pp. 9–34). Clevedon, England: Multilingual Matters.
- Pinter, A. (2005). Task repetition with 10-year-old children. In Edwards, C. and J. Willis. (eds.), *Teachers Exploring Tasks in English Language Teaching*. New York, NY: Palgrave Macmillan.
- Platt, J. T., & Weber, H. (1980). *English in Singapore and Malaysia*. Petaling Jaya, Malaysia: Oxford University Press.
- Poor command of English key reason for graduates' unemployment — director. (2012, June 19). *The Borneo Post*. Retrieved from <http://www.theborneopost.com/2012/06/19/poor-command-of-english-key-reason-for-graduates-unemployment-director/>
- Potter, J. (1996). *Representing reality*. London, England : Sage.
- Potter, M. C., & Faulconer, B. A. (1975). Time to understand pictures and words. *Nature*, 253, 437-438.
- Rahmah Hashim. (1986). Media penyiaran di Malaysia dan hubungannya dengan wanita. *Jurnal Komunikasi*, 2, 34-48.

- Ramiza Darmi , & Albion, P. (2013, July). *Malaysian graduates : What and why*. Paper presented at the 3rd Malaysian Postgraduate Conference, New South Wales, Australia. Retrieved from http://www.academia.edu/4275243/Malaysian_graduates_What_and_why
- Rea, P. (1997). *Designing and Conducting Survey Research*, San Francisco, CA: Jossey-Bass.
- Read, J. (1997). Vocabulary and testing. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 303-320). Cambridge, England: Cambridge University Press.
- Reza Raissi & Fazirah Mohd Nor. (2013, March). *Teachers' perceptions and challenges regarding the implementation of communicative language teaching (clt) in Malaysian secondary schools*. Paper presented at Proceeding of the Global Summit on Education 2013, Kuala Lumpur. Retrieved from http://worldconferences.net/proceedings/gse2013/papers_gse2013/236%20Reza%20Raissi%20&%20Fazirah%20bt%20Mohd%20Nor.pdf
- Rosa, E., & O'Neill, M. D. (1999). Explicitness, intake, and the issue of awareness. *Studies in Second Language Acquisition*, 21, 511–556.
- Rosenthal, R., & Rosnow, R. L. (1991). *Essentials of behavioral research: Methods and data analysis* (2nd ed.). New York, NY: McGraw-Hill.
- Rost, M. (1990). *Listening in language learning*. London, England: Longman.
- Saadiah Darus and Kaladevi Subramaniam. (2009). Error Analysis of the Written English Essays of Secondary School Students in Malaysia: A Case Study. *European Journal of Social Sciences*. 8 (3), 483-495.
- Saadiah Darus. 2010. The Current Situation and Issues of the Teaching of English in Malaysia. *Ritsumeikan Studies in Language and Culture*. 22 (1), 19-27.

- Saal, R. E., Downey, R. G., & Lahey, M. A. (1980). Rating the ratings: Assessing the psychometric quality of rating data. *Psychological Bulletin*, 88, 413-428.
- Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, 50, 696-735.
- Samuel, R., & Bakar, D. (2008). The effectiveness of 'VELT' in promoting English language communication skills: A case study in Malaysia. *International Journal of Education and Development using ICT*, 4(3). Retrieved from <http://ijedict.dec.uwi.edu/viewarticle.php?id=559>.
- Sato, C. (1986). Conversation and interlanguage development: Rethinking the connection. In R. Day (Ed.), *Talking to learn: Conversation in second language acquisition* (pp. 5–22). Rowley, MA: Newbury House.
- Scarcella, R.C., & Higa, C. A. (1981). Input, negotiation, and age differences in second language acquisition. *Language Learning*, 31 (2), 409-437.
- Scarcella, R., & Zimmerman, C. (1998). ESL student performance on a text of academic lexicon. *Studies in Second Language Acquisition*, 20 (1), 27-49.
- Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. Basingstoke, England: Palgrave Macmillan.
- Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11, 129–158.
- Schmidt, R. (1993). Awareness and second language acquisition. *Annual Review of Applied Linguistics*, 13, 206–226.
- Schmidt, R. (1994). Deconstructing consciousness in search of useful definitions for applied linguistics. *AILA Review*, 11, 11–26.

- Schmidt, R. (1995). Consciousness and foreign language learning: A tutorial on the role of attention and awareness in learning. In R. W. Schmidt (Ed.), *Attention and awareness in foreign language learning* (Tech. Rep. No. 9). Honolulu, HI: University of Hawai'i, Second Language Teaching and Curriculum Center.
- Schmidt, R., & Frota, S. (1986). Developing basic conversation ability in a second language: A case study of an adult learner of Portuguese. In R. R. Day (Ed.), *Talking to learn: Conversation in second language acquisition* (pp. 237-326). Rowley, MA: Newbury House.
- Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, 52 (3-4), 591-611.
- Sharwood Smith, M. (1987). Comprehension vs. acquisition: Two ways of processing input. *Applied linguistics*, 7, 239-256.
- Shintani, N. (2011). Input-based tasks and the acquisition of vocabulary and grammar: A process-product study. *Language Teaching Research*, 16 (2), 253- 279.
- Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, 86 (2), 420-428.
- Shuttleworth, M. (2009). *Research bias*. Retrieved from <https://explorable.com/research-bias>
- Skinner, B.F. (1974). *About behaviorism*. New York, NY: Vintage Books.
- Snodgrass, J. G., Wasser, B., Finkelstein, M., & Goldberg, L. B. (1974). On the fate of visual and verbal memory codes for pictures and words: Evidence for a dual coding mechanism in recognition memory. *Journal of Verbal Learning & Verbal Behavior*, 13, 27-37.
- Stahl, K. A. D., & Bravo, M. (2010). Contemporary classroom vocabulary assessment for content areas. *Reading Teacher*, 63, 566-578.

- Stevick, E. (1976). *Memory, meaning and method*. Rowley, MA: Newbury House.
- Streiner, D. L. (2003). Diagnosing tests: Using and misusing diagnostic and screening tests. *Journal of Personality Assessment*, 81, 209-219.
- Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 235–253). Rowley, MA: Newbury House.
- Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), *Principle and practice in applied linguistics: Studies in honor of William E. Rutherford* (pp. 125–144). Oxford, England: Oxford University Press.
- Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive processing they generate: A step towards second language learning. *Applied Linguistics*, 16, 371–391.
- Swain, M., & Lapkin, S. (1998). Interaction and second language learning: Two adolescent French immersion students working together. *Modern Language Journal*, 82, 320–337.
- Taglieber, L. K., Johnson, L.L. & Yarbrough, D. B. (1988). Effects of pre-reading activities on EFL reading by Brazilian college students. *TESOL Quarterly*, 22 (3), 455-472.
- Thorndike, R. L. (1971). Concepts of culture fairness. *Journal of Educational Measurement*, 8, 64.
- Thurmond, A. V. (2001). The point of triangulation. *Journal of Nursing Scholarship*, 33(3), 253-258.
- Trochim, W.M.K. (2006) Introduction to Evaluation. Retrieved from <http://www.socialresearchmethods.net/kb/intreval.htm>

- Van de Branden, K. (1997). Effects of negotiation on language learners' output. *Language Learning*, 47, 289-636.
- Van Lier, L. (1996). *Interaction in the Language Curriculum: Awareness, Autonomy, and Authenticity*. London, England: Longman.
- Varonis, E. M., & Gass, S. (1985). Nonnative/nonnative conversations: A model for negotiation of meaning. *Applied Linguistics*, 6, 71-90.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Wagner-Gough, J., & Hatch, E. (1975). The importance of input data in second language acquisition studies. *Language Learning*, 25, 297-307.
- White, L. (1987). Against comprehensible input: The input hypothesis and development of second language competence. *Applied Linguistics*, 6, 95-110.
- White, L. (1991). Adverb placement in second language acquisition: Some effects of positive and negative evidence in the classroom. *Second Language Research*, 7, 133-161.
- West, M. (1953). A general service list of English words. London, England: Longman, Green & Co.
- Wesche, M. B., & Paribakht, T. (1996). Assessing second language vocabulary. *The Canadian Modern Language Review*, 53, 13-41.
- World Bank (2009), *Malaysia productivity and investment climate assessment Update*. Washington, D.C.
- Youngman, M. B. (1979). Assessing behavioural adjustment to school. *British Journal of Educational Psychology*, 49 (2), 258-264.
- Zimmerman, C. (1997). Do reading and interactive vocabulary instruction make a difference? An empirical study. *TESOL Quarterly*, 31(1), 121- 140.