

THE EFFECTS OF VOCABULARY INSTRUCTION

ON ENGLISH LANGUAGE LEARNERS:

A META-ANALYSIS

by

MIJIN WON, M.A.

A DISSERTATION

IN

CURRICULUM & INSTRUCTION

Submitted to the Graduate Faculty  
of Texas Tech University in  
Partial Fulfillment of  
The requirements for  
The Degree of

DOCTOR OF PHILOSOPHY

Approved

Carole Janisch  
Co-Chairperson of the committee

Hansel Burley  
Co-Chairperson of the committee

Amma Akrofi

Accepted

Fred Hartmeister  
Dean of the Graduate School

May, 2008

Copyright 2008, Mijin Won

## ACKNOWLEDGEMENTS

I owe a great many thanks to so many people who have supported me all the way throughout my study to this final achievement.

It is with deep gratitude that I wish to thank Dr. Janisch for her support and friendship that have made it possible for me as an international student to finish my degree at Texas Tech University. Her personality and commitment have inspired me and will continue to motivate me to strive higher in my academic life. Without her kindness, humor, and personal assistance as well as her academic advising, my dream would not have come true.

I owe a tremendous debt of gratitude to my co-chair, Dr. Burley for his guidance and insightful advice on explaining statistical intricacies and procedures. I wish to thank him for his persistent encouragement and the hours he put in reading and consulting my paper. Without his professional knowledge about the study method, this study would not have possible.

I gratefully wish to thank Dr. Akrofi for her encouraging advice from different perspectives that has broadened my vision. Her consistent support and encouragement contributed very much to my confidence in the completion of this study.

I would also like to extend my special appreciation to Dr. Myers and her husband. Dr. Myers has supported me not only in my academic life, but also in my personal life by giving me a graduate assistant job during the summer as well as looking into my personal life as a neighbor. My thanks also go to Mr. Myers. Without his voluntary help, I would have had a much more difficult time in finishing my work.

I also wish to acknowledge all the faculty members of the department of curriculum and instruction: Dr. Peggy Johnson for her cheer and assistance, and Dr. Sally McMillan for her kind comments on my works.

I would like to extend my deep appreciation to all the staff members of the Interlibrary Loan Office. Without their help, this study would have been impossible.

And last, to my family, words are not enough to express my gratitude. I am grateful to my father who is in heaven and my mother. They have believed in me and always encouraged me. To my husband, Dr. Samsup Jo, I am forever indebted. Without him, I could not have earned a PhD in the United States. His dedication and sacrifices helped me make the right decisions through the ups and downs of my life. I wish to dedicate this volume as a testimony of my gratitude to my son, Jaekwon, and my daughter, Harim for their understanding, cheering, and empowering me throughout these years.

## TABLE OF CONTENTS

ACKNOWLEDGEMENT .....	ii
ABSTRACT.....	vii
LIST OF TABLES.....	x
LIST OF FIGURES.....	xi
CHAPTER I: INTRODUCTION .....	1
Summary of the Problem.....	2
Significance of the Study.....	5
Purpose of the Study.....	6
Research Questions.....	7
Discussion of the Questions.....	7
Vocabulary Knowledge.....	9
Principles of Vocabulary Instruction.....	11
Types of Vocabulary Instruction.....	13
Limitations of the Study.....	16
CHAPTER II: REVIEW OF THE LITERATURE.....	19
Introduction.....	19
History of Vocabulary Research.....	20
Characteristics of Current Vocabulary Research.....	25
Characteristics of ESL Vocabulary Research.....	26
Characteristics of EFL Vocabulary Research.....	29
Review of Research on Vocabulary Instruction.....	32
Summary.....	34

CHAPTER III: METHODOLOGY.....	35
Background on Meta-analysis.....	35
Literature Review of Meta-analysis.....	37
Meta-analysis as a Quantitative Study.....	39
Advantage of Meta-analysis.....	40
Procedure.....	41
The literature search .....	42
Selection of the studies.....	42
Coding of the studies.....	43
Measuring study findings.....	47
Analysis.....	50
Pilot Study.....	52
Summary.....	54
CHAPTER IV: PRESENTATION OF RESULTS.....	55
Description of the Study.....	55
Research Publication.....	56
Research Setting.....	56
Learner Characteristics.....	57
Research Design.....	59
Meta-analysis of Study Findings.....	60
Overall Effect of Vocabulary Instruction.....	60
Effects of Instructional Approaches.....	65

Effects of Learning Conditions (ESL vs. EFL).....	69
Effects of Participants Age.....	72
Effects of Language Level.....	76
Effects of Multimedia.....	78
Effects of First Language Supports.....	81
CHAPTER V: DISCUSSION, CONCLUSION AND RECOMMENDATION	
	.....88
Summary of Main Findings.....	88
Discussion and Conclusions.....	92
Effects of Instructional Approaches .....	92
Effects of Learning Conditions .....	94
Effects of Multimedia Use .....	94
Effects of First Language .....	95
Implications for Classroom Teachers .....	96
Limitations of the Study .....	98
Recommendations.....	99
REFERENCE.....	101
APPENDIX .....	109

## ABSTRACT

The effectiveness of vocabulary instruction for first language learners has been reviewed in several studies (Petty, Herold, & Stoll, 1968; Stahl & Fairbanks, 1986; National Reading Panel, 2000). However, the reviews of vocabulary instruction research focusing on learners of English as a second (ESL) or foreign language (EFL) are fewer in number. Because empirical research has explored more fully which vocabulary instruction is more useful for learning English vocabulary, there is a need for a study reviewing the results of research related to ESL or EFL students. Many empirical studies measured the effectiveness of some instruction individually in particular cases. Therefore, each instructional approach needs to be examined, reviewed, and integrated to determine the effectiveness as a whole. The findings will have implications for teaching vocabulary to ESL or EFL students.

With the above purpose in mind, the following research questions guided my study. Each will be answered by integrating previous studies in this meta-analysis.

1. How can vocabulary instruction for English language learners used in quasi-experimental or experimental studies be categorized?
2. To what degree are these instructional methods effective? What are the effects of such methods on learning vocabulary for English language learners' achievement?
3. Under what conditions are these kinds of instruction effective? That is, what are the mediating effects of substantive variables affecting vocabulary instruction, such as



the age of samples, their levels of English, geographical location, and the publication source of study?

To review these various instructional methods with lesser bias and greater validity than the traditional review, a meta-analysis was conducted. Using meta-analysis techniques, the research relating to vocabulary instruction for English language learners was synthesized. Meta-analysis uses effect size as a common metric for comparing outcomes of each experimental or quasi-experimental study. The meta-analytic process has five basic phases (Glass, 1976): (1) formulating the problem, (2) collecting the data, (3) evaluating the data, (4) synthesizing the data and (5) presenting the findings.

The categorization of vocabulary instructional studies indicated the variety of instructional approaches studied in the past three decades. Studies were categorized as contextualized instruction, semi-contextualized instruction, or decontextualized instruction. The three categories of studies were divided by multimedia and non-multimedia. The category of studies used multimedia as a teaching tool measured computer-related learning word programs. These computer programs were subcategorized as to whether they used first language supports in the programs.

The overall effect size was  $d=.69$  for this meta-analysis. Since the effect size is the same as a **z** score, an effect size  $.7$  is the same as  $.2580$  (about 26%). This means that if the control groups were to receive the treatment, their scores should improve about 26% on the average. Even though all of the instruction is effective for English language learners, decontextualizing instruction is the most effective for them. Decontextualizing instruction is focusing on the word meaning separately from context with a flash card or

word list, or any other activities that focus on forms or core meanings of words. Contextual instruction shows the lowest effect size than the other two types of instruction.

The study indicates there are no different effects of vocabulary instruction on learning condition between ESL and EFL. The use of multimedia in vocabulary instruction was found to be more effective than control groups, but the effect size of multimedia groups was lower than the group without multimedia use. Therefore, it could be said that the use of a computer is not better than a classroom teacher. Although there are debates on the effectiveness of first language use for second language learners, the results of this study yielded no differences with or without first language supports, unless the instruction was associated with the other specific factors.

This meta-analysis confirms that vocabulary instruction for ESL or EFL learners has some similarities and dissimilarities to that of first language learners. To determine more effective instruction for EFL or ESL learners, a consideration of the characteristics of learners and environments as well as instructional approaches should be made. From this study, it was concluded that effective instruction for English language learners depends on the conditions of vocabulary instruction. ESL and EFL classroom teachers need to consider the difference depends on the conditions of vocabulary instruction. ESL and EFL classroom teachers need to consider the differences of first and second language vocabulary acquisition as well as student learner characteristics.

## LIST OF TABLES

3.1 Coding Categories Used In Analysis.....	44
4.1 Learners Characteristics.....	58
4.2 Overall Effect Size Mean for the 43 Studies of Vocabulary Instruction .....	61

## LIST OF FIGURES

4.1	Displays of effect sizes from studies of vocabulary instruction.....	63
4.2	Effects of instructional approaches.....	68
4.3	Effects of learning conditions.....	70
4.4	Effects of instructional approaches on leaning condition.....	71
4.5	Effects of participant age.....	73
4.6	Effects of instructional approaches for college students.....	75
4.7	Effects of the level of language proficiency.....	77
4.8	Effects of multimedia use.....	79
4.9	Effects of multimedia use on EFL condition.....	80
4.10	Effects of first language supports.....	82
4.11	Effects of first language supports on semi-contextualized instruction.....	84
4.12	Effects of first language supports on contextualized instruction.....	85
4.13	Effects of first language supports on multimedia use.....	87

## **CHAPTER I**

### **INTRODUCTION**

As a second-language learner of English and as a doctoral student using English as the sole language for my academic work, I recognize that vocabulary has been very important for speaking, listening, and studying. I have also come to appreciate the fact that acquiring new vocabulary is not easy, especially when studying for the TOEFL or Graduate Record Examination. For those exams in particular, my study method consisted of rote memorization. However, I now believe there might be a more effective way to acquire new vocabulary. Folse (2004) reported that the college students of English as Second Language (ESL) in intensive academic programs expressed a strong desire for vocabulary instruction. When students were asked what could improve ESL programs, they ranked “more vocabulary instruction” as the most important improvement.

The National Reading Panel (NRP) examined more than 20,000 research citations to determine how vocabulary can be taught best and related to the reading comprehension process (2000). However, citations were removed from the report if they dealt exclusively with disabled learners or other special populations, including second-language learners. Therefore, there is a need for studies reviewing research excluded from the report of the NRP. Thus, this study will focus on the research which dealt with English language learners to elicit the instruction effective for them. In comparing the results of effective vocabulary instruction from the NRP’s report concerning native speakers of English with effective vocabulary instruction concerning for English

language learners, this study explores the effectiveness of vocabulary instruction for the English language learners.

I believe that one specific instruction does not have same effects for different characteristics of English language learners, e.g., beginning, intermediated, and advanced learners or younger and older learners. It seems to be impossible to argue which type of instruction was better than the others without knowing characteristics of learners and learning conditions. However, the differences in results of vocabulary studies strongly suggest the possibility that some vocabulary instruction may be more effective than others for English language learners in some cases. Therefore, the research needs to determine which instruction is most effective for learners of different ages, proficiency levels, language groups, and social setting even though it is true that incidental vocabulary learning is possible. The purpose of this research is to investigate the effects of vocabulary instruction on acquiring English vocabulary with English language learners of all ages, proficiency levels, and academic settings. These findings could help ESL or English Foreign Language (EFL) classroom teachers in the US or other countries in the world to develop their curricula of vocabulary instruction that support students' effective learning.

### **Summary of the Problem**

How learners of ESL or EFL can acquire a new vocabulary effectively has been studied following various theoretical and empirical frameworks of language acquisition. One of the major debates in vocabulary learning is whether incidental vocabulary

learning is better than explicit instruction. It is believed that vocabulary is better learned incidentally for both first and second language learners. Krashen (1989) concluded after reviewing 144 studies that incidental vocabulary learning gives better results than intentional vocabulary learning. Huckin and Coady (1999) also concluded that vocabulary learning predominantly occurs incidentally through reading.

However, many second language vocabulary researchers are concerned that natural or incidental vocabulary acquisition is simply not efficient enough to produce the desired rates of learning. Natural context is not an especially rich source of information about word meanings. According to Nagy and Scott (2000), free reading is the least effective way to reach the goals if there are particular words a student wants to learn. Nation (1990), and Coady, Magoto, Hubbard, Graney, and Mokhtari (1993) suggested that a second language learner needs to know approximately 2,000 high-frequency words to understand about 85% of most texts. They argued that direct instruction of those words allowed students to learn the high-frequency words that they needed to understand texts. Therefore, the direct instruction of second language vocabulary should be mandatory for students at the beginning level, and would be beneficial for all levels of second language learners.

Chall's work (1987) suggested that decisions to use incidental versus conscious approaches can only be made by considering students' ages and proficiency levels. Coady (1993) also concluded after exploring the basic argument for a mixed approach to vocabulary acquisition in ESL that the basic or core vocabulary should be taught, but less frequent vocabulary will be learned 'naturally' via context, but even in that case, techniques for that purpose should be taught. Carter and McCarthy (1988) concluded that

a mixture of approaches should be adopted since there are advantages and disadvantages between context-based inferential strategies and some other explicit vocabulary learning approaches such as key-word techniques, translation in pairs, using a monolingual or bilingual dictionary. According to Sternberg (1987), if most vocabulary is learned from context, one should not conclude that this “is the fastest or most efficient way of learning specific vocabulary.” Several studies have found positive evidence supporting the use of explicit vocabulary instruction in conjunction with extensive reading (Paribakht & Wesche, 1999). In fact, language teachers use a variety of methods to teach vocabulary as well as good learners use a more diverse set of vocabulary-learning strategies. In summary, in spite of the evident role of reading in much advanced vocabulary instruction, the incidental learning process is slow, often misguided, and seemingly haphazard, with differential outcomes for different learners, word types, and context.

Stahl and Fairbanks (1986) conducted a meta-analysis on the effectiveness of vocabulary instruction. Their research was one of the first 100 references of meta-analysis studies to use meta-analysis in order to consolidate and review an issue of importance to elementary education (Nagy & Scott, 2006) and has become a classic in the field of vocabulary instruction research. They pointed out the needs of meta-analysis, because there were mixed results of studies about effective instruction, and they premised that some methods of vocabulary instruction may be more effective than others. Given the diversity of research findings, meta-analytic procedures appeared to be a reasonable approach to this area of study.

After Stahl and Fairbanks’ study, there were some meta-analysis studies relative to vocabulary instruction for first language learners (Fukkink & de Glopper, 1998;



Swanborn & de Glopper, 1999) or some meta-analysis studies relative to reading or cognitive skills for English language learners (Wa-Mbaleka, 2006); however, none of the studies reviewed within the meta-analysis focused only on vocabulary instruction for ESL and EFL learners. Therefore, my meta-analysis builds on the previous literature by analyzing data from research in the past three decades and from studies that were not included in NPR reports, i.e., studies of ESL and EFL learners. In the previous research included in my meta-analysis, the diversity of research methods, different sample (from children to adults, from beginners to advanced learners, and from living in English speaking country to the other countries which speak other first languages) and different criterion measures existed. Nonetheless, even with these kinds of differences, by employing the meta-analysis techniques, it was possible to explore the effects of vocabulary instruction.

### **Significance of the Study**

The effectiveness of vocabulary instruction for first language learners has been reviewed in several studies (Petty, Herold, & Stoll, 1968; Stahl & Fairbanks, 1986; National Reading Panel, 2000). However, the reviews of vocabulary instruction research focusing on ESL and EFL learners are fewer in number. Because empirical research has explored more fully which vocabulary instruction is more useful for learning English vocabulary, there is a need for a study reviewing the results of research related to ESL or EFL students. It might not be much different from the results concerning the first language learners. However, this study focuses on vocabulary instruction research for

English language learners because intermediate and advanced level English learners at the college level consistently point out their lack of vocabulary as a major obstacle to achieving English proficiency (Folse, 2004). Since not much is known about the effectiveness of the specific vocabulary instruction in relation to the purpose of learning English, much more research is needed concerning the appropriate vocabulary instruction for English language learners who are taking English classes as a second language or foreign language in colleges and universities as well as elementary and secondary schools. Many empirical studies measured the effectiveness of some instruction for particular cases. Therefore, each instruction needs to be examined, reviewed, and integrated to determine the effectiveness as a whole.

### **Purpose of the Study**

The purpose of this study is to draw generalizations about vocabulary instruction for English language learners by doing the following: (1) categorizing the kinds of vocabulary instruction or instructional methods used in the previous research, (2) comparing the effectiveness of the instruction based on the achievement outcomes, and (3) determining the effective conditions for English language learners including ESL and EFL learners of all proficiency levels. To review these various instructional methods with lesser bias and greater validity than the traditional review, a meta-analysis was conducted. Using meta-analysis techniques, the research relating to vocabulary instruction for English language learners was synthesized. Meta-analysis uses effect size as a common metric for comparing outcomes of each experimental or quasi-experimental

study. In this synthesis, experimental studies that were studied after 1980 were analyzed. By employing the meta-analysis techniques, it was possible to explore the effects of vocabulary instruction for English language learners.

### **Research Questions**

With the above purpose in mind, the following research questions were developed to guide my study. They were answered by integrating previous studies in this meta-analysis.

1. How can vocabulary instruction for English language learners used in quasi-experimental or experimental studies be categorized?

2. To what degree are these methods of instruction effective? What are the effects of such methods on learning vocabulary for English language learners' achievement?

3. Under what conditions are these kinds of instruction effective? That is, what are the mediating effects of substantive variables affecting vocabulary instruction, such as the age of samples, their levels of English, geographical location, and the publication source of study?

### ***Discussion of the Questions***

With these questions, I was interested in the role that effective vocabulary instruction played with English language learners who were trying to learn English mostly for academic purposes. Are some vocabulary instructional methods more effective

than others for English learners if some specific conditions exist? I also wanted to know how English learners of different backgrounds (country of origin, academic disciplines, and proficiency levels) used the vocabulary instruction. I wanted to know how learners of English use the vocabulary instruction differently from the native speakers of English by comparing the results of this study to previous reviews of analysis of vocabulary instruction for first language learners as well. Finally, I was interested in learning what kinds of vocabulary instruction help both ESL and EFL students effective acquisition of new words in different learning settings with the relation of different learners' characteristics.

In the process of framing the questions for this study, I explored the meanings of the following terms and their use in previous research:

#### English as a Second Language (ESL)

English studied by non-native speakers as a second language in an English-speaking country, as in the United States.

#### English as a Foreign Language (EFL)

English studies by non-native speakers as a foreign language in an environment where English is not spoken as the first language.

#### English Language Learners (ELL)

Individuals who are in the process of learning English after learning a first or other language.

## Vocabulary Knowledge

The study by Nagy and Scott (2000) focused on how children learn the meaning of new words. Children's vocabulary growth must be based on recognition of the complexity of word knowledge. They pointed out five characteristics of vocabulary knowledge.

First, knowing a word is a matter of degree, not all-or nothing. This is known as the incremental view of word knowledge. Dale (1965) proposed four stages of degree of word knowledge; (1) never said it before, (2) heard it but doesn't know what it means, (3) recognizes it in context as having something to do with, and (4) knows it well (see Nagy & Scott, 2000). After that, there are some variations proposed by several researchers. Drum and Konopak (1987) outlined six levels of word knowledge; (1) knows a word orally but not in written form, (2) knows a words meaning but cannot express it, (3) knows a meaning but not the word for it, (4) knows the partial meaning of a word, (5) knows a different meaning for a word, and (6) knows neither the concept nor the word.

Paribakht and Wesche (1993) created a five stage "Vocabulary Knowledge Scale:"

Stage 1: The word is not familiar at all.

Stage 2: The word is familiar but the meaning is not known.

Stage 3: A correct synonym or translation is given.

Stage 4: The word is used with semantic appropriateness in a sentence.

Stage 5: The word is used with semantic appropriateness and grammatical accuracy in a sentence.

This incremental view of word learning helps explain how a great deal of vocabulary knowledge can be gained incidentally from context.

The second characteristic of vocabulary knowledge is referred to as multidimensionality: word knowledge consists of several qualitatively different types of knowledge. Nation (1990) offered eight aspects of word knowledge: (1) the spoken form of a word, (2) the written form of the word, (3) the grammatical behavior of the word, (4) the collocation behavior of the word, (5) the frequency of the word, (6) the stylistic register constraints of the word, (7) the conceptual meaning of the word, and (8) the associations the word has with other related words.

The third characteristic is the idea of polysemy: words often have multiple meanings. Word meanings are flexible, and always nuanced in some way by the context in which they occur. The meaning of a word must be inferred from context.

Fourth is interrelatedness: one's knowledge of any given word is not independent of one's knowledge of other words.

The fifth and last characteristic is heterogeneity: what it means to know a word differs substantially depending on the kind of word.

Although the Nagy and Scott (2000) study on the complexity of vocabulary knowledge was done with first language children, it is highly likely that vocabulary learning by second language learners also consists of many degrees of knowledge. For instructional purposes, Stahl (1983, 1986) proposed two types of knowledge about words: definitional information and contextual information. Definitional information comprises knowledge of the logical relationship between a word and other known words, as in a dictionary definition involving knowing a definition, synonym, antonym, or affixes, and

so on. Contextual information can be defined as knowledge of the core concept of the word and how that concept changes in different contexts. Depending on the student's knowledge of a word, effective vocabulary instruction will employ different techniques. Huckin and Bloch (1993) pointed out that second language readers rely heavily on vocabulary knowledge, and a lack of vocabulary knowledge is the largest obstacle for second language readers to overcome. They found that the main obstacle for second language readers is not lack of reading strategies but rather insufficient vocabulary knowledge in English.

Also, the strong relationship between vocabulary knowledge and academic achievement were researched. Anderson and Freebody (1983) hypothesized that vocabulary knowledge is strongly related to comprehension because (1) understanding words enables readers to understand passages, and/or (2) verbal aptitude underlies both word and passage comprehension, and/or (3) vocabulary knowledge may be related to a person's store of background information.

Based on the above explanation, the term "vocabulary knowledge" is used in this study to mean the every dimension of complex word knowledge related to comprehension.

### Principles of Vocabulary Instruction

There are several factors related to effective vocabulary learning for first language learners as well as English language learners as a second language. Beck and Mckeown (1991) found that (1) direct vocabulary instruction can increase comprehension if the text contains the words taught, (2) vocabulary instruction needs to be extensive and include

frequent encounters with the word to affect comprehension, (3) instruction in vocabulary should include associating new words outside class, and (4) instruction can be most effective when words are related to each other meaningfully. Reviews of successful instructional practice reveal that direct instruction leads to improved reading comprehension and has the following components: (1) multiple exposures to the word, (2) elaboration and discussion of word meaning connecting with the students' prior knowledge to facilitate deep processing of words and their meaning (definitional and contextual meaning of words), and (3) application of word meanings to new situations (Beck, Perfetti, & McKeown, 1982).

Blachowiz and Fisher (2002) developed four main principles to guide vocabulary instruction based on their review of vocabulary research. Those principles are that students should:

1. Be active in developing an understanding of words and ways to learn them.
2. Personalize word learning.
3. Be immersed in words.
4. Build on multiple sources of information to learn words through repeated exposures.

In the detailed recent reports by the National Reading Panel (2000), they offer the following implications for practice based on findings from first language instruction.

1. Vocabulary should be taught both directly and indirectly.
2. Repetition and multiple exposures to vocabulary items are important.
3. Learning in rich context is valuable for vocabulary learning.
4. Vocabulary tasks should be restructured when necessary.



5. Vocabulary learning should entail active engagement in learning task.
6. Computer technology can be used to help teach vocabulary.
7. Vocabulary can be acquired through incidental learning.
8. How vocabulary is assessed and evaluated can have differential effects on instruction.
9. Dependence on a single vocabulary instruction method will not result in optimal learning.

Barcroft (2004) discussed five principles for effective second language vocabulary instruction with emphasis on lexical input processing: (1) present new words frequently and repeatedly in the input, (2) use meaning-bearing comprehensible input when presenting new words, (3) limit forced output during the early stages of learning new words, (4) limit forced semantic elaboration during the initial stages of learning new words, and (5) progress from less demanding to more demanding vocabulary-related activities. He emphasized provision of new words in the input and incremental development of word knowledge and drew attentions to how learners must allocate their limited mental resources in order to acquire multiple components of words knowledge, including word form, form-meaning mapping, and second language specific usage.

### Types of Vocabulary Instruction

The debate on effective vocabulary learning concerning incidental acquisition affect the different types of vocabulary instruction according to their focus on either context or definition. Depending on how much emphasis is placed on context or definition, five types of vocabulary instruction are used in practice.

1. **Definitional only:** The only information provided is a definition, synonym, and so forth. There are no examples of the word used in context.
2. **Definitional emphasis:** Some exposure is given to the word in context, but emphasis is on the child learning the definition.
3. **Balanced:** A balance or near balance between definitional and contextual information is given,
4. **Contextual emphasis:** Although a definition is given, the major emphasis is on learning the word in context.
5. **Context only:** The child is exposed only to each word in context, with no attempt to have the child derive a definition.

Oxford and Crookall (1990) described a number of techniques by which vocabulary instruction has been handled, mishandled, or avoided almost entirely by second language teachers. They divided vocabulary learning techniques into four categories: decontextualizing, semi-contextualizing, fully contextualizing, and adaptable.

1. **Decontextualizing techniques:** remove the word as completely as possible from any communicative context (word lists, flashcards, dictionary use).
2. **Semi-contextualizing techniques:** allow some degree of context but fall short of full contextuality (word groupings, word or concept association, visual imagery, aural imagery, keyword, physical response, physical sensation, semantic mapping).
3. **Fully contextualizing techniques:** embed the new words in a more or less normal communicative context (reading and listening practice, speaking and writing practice).

4. Adaptable techniques (structured reviewing): reinforce other techniques at any part of the contextuality continuum. The technique entails going back over second language vocabulary at different intervals, at first close together and then increasingly far apart. Structured reviewing is scientifically based on memory principles, which highlight the importance of primacy, recency, duration, spacing, pacing, and linking. Second language textbooks typically fail to make any overt suggestions to learners about these techniques.

Conceptually, vocabulary can be measured in many ways. One major distinction in the measurement of vocabulary parallels the receptive and productive distinction. Another distinction is made between reading and writing vocabulary, and speaking and listening vocabulary. There are so many definitions of vocabulary; accessing vocabulary has many variables in vocabulary research. Since many studies examined involve unique instruction, the National Reading Panel attempted to categorize the following methods: (1) explicit instruction, (2) indirect instruction, (3) multimedia methods, (4) capacity methods, and (5) association methods.

1. Explicit instruction: students are given definitions or other attributes of words to be learned (pre-teaching of vocabulary, analysis of word roots or affixes).
2. Indirect instruction: students are exposed to words or given opportunities to do a great deal of reading (wide reading to increase vocabulary).
3. Multimedia methods: vocabulary is taught by going beyond text to include other media (semantic mapping, graphic representation).
4. Capacity methods: these methods attempt to reduce the cognitive capacity devoted to other reading activities by practicing to make them more nearly

automatic (concentrate on meaning of words rather than their orthographic or oral representation).

5. Associate methods: learners are encouraged to draw connections between what they do know and word they encounter that they do not know (associations are semantic or contextual, or imagery students invoke in learning the words).

Contextualizing techniques by Oxford and Crookall (1990) are close to indirect instruction in the NRP category while decontextualizing to explicit instruction.

Association methods by NRP (2000) are definitely one of semi-contextualizing techniques. Multimedia methods are similar to semi-contextualizing techniques except they use multimedia as a tool. There are some similarities between adaptable techniques and capacity methods, which focus on the process of learning. Even though there are some overlaps with each other in Oxford and Crookall's study (1990) and the NRP's reports, the basic categories were divided by distinguishing between context and definition, or explicit and implicit. After distinguishing between two levels, some other categories of instruction e.g., semi-contextualizing, or capacity methods are added in order to cover every instruction used in learning or teaching words. Based on the above explanation of "types of vocabulary instruction," the current study attempted to categorize the types of vocabulary instruction for ESL and EFL learners.

### **Limitations of the Study**

One of the limitations may be that this research was conducted on the studies collected by the researcher. I attempted to gather as many studies as possible,

unavailability and prohibitive costs prevented me from accessing some of the studies relevant to the entire collection for the population of the study.

Another limitation could be the accuracy of coding. This study used only one coder; therefore it was possible to code with a researcher's bias in the coding of each selected study. Even though I conducted a reliability check of all of the studies for accuracy, there are no intercoder checks which are suggested in coding for reliability.

Also, this study was limited by the characteristics of a meta-analysis, criticized from other perspectives of research as "mixing apples and oranges." Alexander, Schallert, and Hare (1991) point out that reviews of vocabulary research appear inconclusive, conflicting, or inapplicable because researchers are quite literally not talking about the same issues, often using the same term in considerably different ways. Definitions of independent and dependent variables, subjects, instruments, and experimental designs vary from study to study. For examples, the definition of knowing a word depends on how a researcher decides acquiring a word meaning. The Vocabulary Knowledge Scale by Paribakht and Wesche (1993) considered five stages of vocabulary acquisition from stage one (the word is not familiar at all) to stage five (the word is used with semantic appropriateness and grammatical accuracy in a sentence). Which of the five stages actually explains the acquisition of a word was not clear and was different from study to study that I used in this meta-analysis.

In addition to the definition of terms used in each study, this study used a wide range of studies. Weak studies might lower the validity and reliability of their reported outcomes. The rationale for including those studies will be discussed in Chapter III.

Finally, my relatively limited research experience might be thought of as another limitation of this study, even though I have conducted a meta-analysis study before. Better interpretation skills for emerging issues would enhance this study. Other limitations to this study include my biases and personal assumptions as mentioned above. As a researcher conducting a meta-analysis by myself, I had to keep in mind the reliability and validity of my study.

## **CHAPTER II**

### **REVIEW OF THE LITERATURE**

#### **Introduction**

In relation to vocabulary instruction research and English learners from different backgrounds, I investigated prior research studies before entering the research site, and continued to consult this research after I collected data. Vocabulary instruction research provides implications for classroom teachers and supports background theory for curriculum. A review of literature reveals that vocabulary instruction varies from study to study. Results of the studies are rich and diverse depending on the purpose of the study. These inconclusive results need a synthesized study of vocabulary instruction as a whole.

Before beginning the meta-analysis, I reviewed the literature that (1) described the history of vocabulary instruction for English language learners, (2) described current characteristics of vocabulary instruction research, and (3) reviewed the research that has most affected vocabulary instruction studies. This review provided a starting point for categorizing the kinds of vocabulary instruction, determining the effectiveness of each instruction, and finding mediating reasons for effectiveness or ineffectiveness.

First, the history in this review provides vocabulary instruction developed with relation to literacy theory and presents characteristics of background knowledge for each type of instruction. Then I review the selected reports that describe the current characteristics of vocabulary instruction research. Finally, this chapter ends with a

discussion of vocabulary instruction research and the unresolved debates about English learners.

### **History of Vocabulary Research**

The history of research on vocabulary instruction is complex. Until 1950, vocabulary research focused on four areas: (1) vocabulary size at various ages, (2) the relationship between vocabulary and intelligence, (3) identifying the most useful words to know, and (4) identifying a core of words that make text more understandable ( Irvin, 1990). However, the study of vocabulary was one of the weaker areas in early research for English language learners while the study about grammatical and phonological structure had been dominant throughout the 1940s, 1950s, and 1960s. Charles Fries' *Teaching and Learning English as a Foreign language* (1945, as cited in DeCarrico 2001) was the most influential study for this tradition based on behaviorist psychology. It valued audio-lingual method as a good way to learn second languages by paying systematic attention to intensive drills of basic sentence patterns and their pronunciation (DeCarrico, 2001). The basic assumption was that once students learned the structural frames, lexical items could be learned later to fill the grammatical slot in the frames. Direct method or audio-lingual method in this period emphasized oral skills, accurate production, and limited vocabulary knowledge as a way to build good language use habits. From this perspective, good language habits would eventually lead to an increased vocabulary.



Revolutionary changes in linguistic theory were brought by Chomsky (1957). In his work, language teaching was viewed as a rationalist's framework rather than the behaviorists' notion of habit formation. The central assumption was that language is represented as a speaker's mental grammar, in other words, a set of abstract rules for generating grammatical sentences. Since, language learning was considered as rule acquisition, not habit formation, vocabulary was somewhat important; however, rule learning still has a place in language learning (Decarrico, 2001).

In 1970's, Hymes's concept of communicative competence emphasized the sociolinguistic and pragmatic factors governing effective use of language (Hymes, 1972). He was especially concerned about using language for meaningful communication, including the appropriate use of language in particular social contexts. Since the communicative language teaching promoted fluency over accuracy, lexical competence was a central part of communicative competence. In other words, teaching vocabulary started to become a central part of teaching language in contrast to early language research. However, during the 1970s contemporary linguistics and cognitive psychology supported the psycholinguistic approach, which focused on guessing the meaning of unknown words through the use of contextual clues (Coady, 1993). Psycholinguistic studies provided insights concerning mental processes involved in vocabulary learning, such as memory, storage, and retrieval.

In the 1980s there seems to have developed a reaction against the psycholinguistic model. The research trend in this period emphasized the role of lexis in large units of language beyond the single word-form. The meaning has to be reinterpreted constantly throughout a text because of the interaction of a number of text features such as lexical

cohesion, subordinators, pragmatic consideration, coherence relations, and genre structures (Coady, 1993). This interactive approach argues that the proficient readers utilize both bottom-up and top-down processing, and that successful comprehension is the result of an interaction between both types of processing. Within this approach, schema theory emphasizes the role of preexisting knowledge which the learner relates to the input from the text interactively. Thus, interactional activities in this framework emphasize teaching students to take advantage of all of their prior knowledge. As a result, vocabulary acquisition is viewed in terms of the students' background knowledge of concepts as well as of word forms.

In addition, in the 70's and 80s' the communicative approach and interactional approach focused on implicit, incidental learning. Incidental vocabulary learning is defined as learning that occurs when the mind is focused elsewhere, such as on understanding a text or using language for communicative purpose. In a review of 144 studies, Krashen (1989) argued that incidental acquisition of vocabulary occurs through the operation of his Input Hypothesis, which proposes that learners acquire a second language when they are exposed to comprehensible input.

A number of studies which support this hypothesis have shown that guessing from context can lead to vocabulary acquisition. Raptis (1997) showed that many current second language reading textbooks promote the assumption that vocabulary is best learned incidentally by guessing from context. Based on this learning theory, teachers encouraged students to recognize clues to word meanings in context and to use monolingual dictionaries rather than bilingual dictionaries, and textbooks emphasized inferring word meaning from context.

In the review of incidental vocabulary learning, Huckin and Coady (1999) stated some advantages of incidental vocabulary learning over direct introduction: (1) It is contextualized, giving the learner a paired-associate exercises, (2) it is pedagogically efficient in that it enables two activities--vocabulary acquisition and reading--to occur at the same time, and (3) it is more individualized and learner-based because the vocabulary being acquired is dependent on the learner's own selection of reading materials.

However, Huckin and Coady in the same article point out some limitations of incidental learning: (1) guessing is imprecise because many reading tasks call for precise interpretation, (2) accurate guessing require accurate word recognition and careful monitoring because there are many deceptive lexical items that can easily mislead the learner, (3) guessing takes time and thus slows down the reading process, (4) guessing is effective only when the context is well understood and almost all of the surrounding words in the text are known, (5) guessing requires good reading strategies, (6) guessing often does not translate into acquisition, and (7) guessing is not effective in the acquisition of multiword lexical items. In spite of the above, they concluded that the incidental learning is still seen as an important part of vocabulary building, especially among advanced learners, but it requires a great deal of prior training in basic vocabulary, word recognition, metacognition, and subject matter. In fact, most scholars agree that, except for the first few thousand most common words, vocabulary learning predominantly occurs through extensive reading incidentally, with the learner guessing at the meaning of unknown words.

However, Hulstijn (1992) reported that the number of new words learned incidentally is relatively small compared to the number of words learned intentionally.

Incidental vocabulary learning tends to be incremental and slow even with the use of a dictionary and the inferring strategy. The study by Hulstijn, Hillander, and Greidanus (1996) pointed out why second language learners could not have enough learning incidentally. The authors suggest the following reasons: (1) learner failed to notice the new words, (2) they noticed the new word, but ignored them, (3) they do not focus their attention on the unknown word, (4) they infer the meaning from context incorrectly, and (5) the low frequency of most unknown prevents effective learning. It was emphasized by Hulstijn (1992) that both incidental and intentional learning should exist together in vocabulary instruction for second language learners.

In fact, Chall's work (1987) suggested that decisions to use incidental versus conscious approaches can only be made by considering students' ages and proficiency levels. Coady (1993) also concluded after exploring the basic argument for a mixed approach to vocabulary acquisition in ESL that the basic or core vocabulary should be taught, but less frequent vocabulary will be learned "naturally" via context, but even in that case, techniques for that purpose should be taught. Carter and McCarthy (1988) concluded that a mixture of approaches should be adopted since there are advantages and disadvantages between context-based inferential strategies and some other explicit vocabulary learning approaches such as key-word techniques, or translation in pairs, or using a monolingual or bilingual dictionary.

In summary, in spite of the evident role of reading in much advanced vocabulary acquisition, there are some problems from the perspective of effective learning. In incidental acquisition through reading, the acquisition process is slow, often misguided, and seemingly haphazard, with differential outcomes for different learners, word types,

and context. According to Sternberg (1987), even if most vocabulary is learned from context, one should not conclude that this is the fastest or most efficient way of learning specific vocabulary.

### **Characteristics of Current Vocabulary Research**

Folse (2004) pointed out eight categories for recent trends in vocabulary research: (1) How many words and which words do learners need to know? (2) How do second language learners' vocabularies develop? (3) Why are some words more difficult to learn than other? (4) Is second language vocabulary learned more easily through natural context or through direct instruction? (5) Which vocabulary learning strategies do students employ? (6) Which types of practice activities promote vocabulary learning? (7) What effect do certain types of marginal glosses and internet annotations have on incidental vocabulary learning? (8) How does using a dictionary impact vocabulary acquisition? All the questions deal with vocabulary instruction from various perspectives, which reveal the characteristics of current vocabulary instruction research.

Much research has been conducted on which kind of instruction works best. However, there is little difference in research trends depending on where research is conducted. I reviewed and divided material into two groups: research conducted in English speaking countries (ESL learning conditions) and research conducted in non-English speaking countries (EFL learning conditions). The characteristics of the ESL group are discussed separately from those in EFL conditions.

### *Characteristics of ESL Vocabulary Research*

For vocabulary instruction research for English language learners under ESL conditions, the studies focus on whether a structured vocabulary approach or some kinds of vocabulary programs are effective. Sanaoui (1995) reported that learners in Canada who have a structured learning approach are more successful in retaining vocabulary taught in their classes. A structured approach is found to be more effective than an unstructured approach for both beginning and advanced learners. However, Lessard-Clouston (1996) concluded that a more structured approach would not necessarily result in more vocabulary learning. Rather, the individual nature of vocabulary learning including a learner's learning style, motivation, previous education, may play an important role. Of course, there is a study that reveals the same result in first language vocabulary acquisition about incidental learning. Paribakht and Wesche (1999) found that that most vocabulary learning occurs naturally when learners attempt to understand new words when they hear or read them in context. But they added later that, reading-based approaches might reasonably be combined with explicit instruction for an initial core of several thousand frequently used words to bring learners to a threshold level for text comprehension (Wesche & Paribakht, 2000).

Gaudio (2003) reported that vocabulary is acquired through the intensive vocabulary building program which included the use of small group vocabulary instruction, computerized vocabulary programs, vocabulary software, vocabulary based games, and mini vocabulary dictionaries. In addition, there are some studies regarding the effectiveness of video clips in teaching unknown vocabulary (Al-Seghayer, 2001), the effectiveness of the collaborative database using online-resources for learners who have

moved beyond the elementary level (Cobb & Horst 2001), and the reevaluation of the dictionary use in the L2 reading class (Fraser, 1999).

The other factor affecting second language vocabulary instruction, not considered in first language acquisition, is the translation effect. Prince (1996) pointed out that the effectiveness of translation learning depends on the learners' proficiency. The use of primary language support is a strategy widely recommended for second language vocabulary instruction; however, the research opposing the use of native language strategies was prevalent as well. Despite controversies surrounding the use of primary language, that strategy can come in many forms, from direct instruction in the native language to translation of worksheets. Fraser's study (1999) shows that consulting a dictionary to confirm inference is a valuable metacognitive strategy for lexical acquisition. Kroll and Curley (1988) stated that ESL students use translation exclusively in the beginning stage compared to students in advanced stages.

Sautermeister (1989) reported vocabulary learning behavior among university, non-specialist learners of English who were consistently presented with new words in context, but who were not satisfied until they found a first language equivalent to assist their learning. This behavior is similar to that described for beginners above. In summary, it appears that learning vocabulary in context is widely perceived by the teaching profession as desirable; however, it is true that the translation condition is sometimes better in the beginning stage.

The process of inferring (Bot, Paribakht, & Wesche, 1999) and the process of negotiation (Smith, 2004) as well as gestures or non-verbal behaviors (Lazaraton, 2004) also affect the effectiveness of instruction in the studies conducted under ESL conditions.

Comparing studies with native English learners, Zareva, Schwanenflgel, and Nikolova, (2005) reported that vocabulary size, word frequency effects, number of associations, and with-group consistency are more effective whereas learners' metacognitive awareness is not proficiency dependent. Also Kojic-Sabo and Lightbrown (1999) investigated the differences between ESL and EFL vocabulary learning. Students in the two settings exhibit some differences with regard to what strategies they used and to what extent. However, vocabulary learning is not necessarily related to those conditions and is rather related to other factors. Extensive strategy use is linked to success in language learning, whereas lack of effort on the learners' part relates to poor achievement. Time and learned independence were two measures most closely related to success in vocabulary learning and higher overall English proficiency.

Much research has been conducted about second language vocabulary knowledge and vocabulary processing compared to first language instruction. Second language learners' mental processing in the target language is different from first language learners (Wolter, 2001). In other words, phonology does not play an important role for second language learners compared to the role of phonology in first language vocabulary learning; rather, semantic factors seem to play more roles in learning second language vocabulary. Quin (2002) researched vocabulary knowledge depth, and argued that the dimension of vocabulary depth is as important as vocabulary size in predicting performance on academic reading for ESL students. Knowledge of word meaning showed a higher likelihood of being remembered more than of it being forgotten (Schmitt, 1998).



In summary, for vocabulary instruction research for English language learners in ESL conditions, research first focuses on whether a structured vocabulary approach or a vocabulary building program is effective; second, whether specific supports, such as online-resources, dictionaries, translations, or technologies, are more effective than incidental learning; third, how second language vocabulary knowledge and lexical processing are different from those of first language.

### ***Characteristics of EFL Vocabulary Research***

The research conducted under EFL conditions showed little difference from research under ESL conditions. EFL students' purpose for learning English is similar to that of ESL students, but their different environments have affected the focus of vocabulary research.

Qian (1996) showed results that learning vocabulary by word lists is more effective than learning vocabulary in context. He did not claim that lexical guessing in context is not effective, but he was concerned about the learner's background, needs, preference, and learning style as important factors for acquiring new words, especially for Asian students. However, Gu and Johnson (1996) conducted a study about Chinese students in China and found contradictions to popular beliefs about Asian learners. The participants did not use memorization; rather, they used more meaning-oriented strategies. Additionally, Fan's study (2003) about Chinese students in Hong Kong showed similar results that guessing unknown words is most often used to learn vocabulary, particularly for the high level vocabulary learners. Laufer and Hill (2000) conducted a study about dictionary use. Their results showed that different people have

different lookup preferences and the use of multiple dictionary information seems to reinforce retention. Also, using first language translation is effective in second language vocabulary learning similar to dictionary use and first – second language paired word lists.

Hulstijn and Laufer (2001) studied the retention of vocabulary, which is related to the amount of task-induced involvement load: retention is highest in the composition tasks, lower in reading plus tasks, and lowest in reading only tasks. Hill and Laufer (2003) also concluded that two form-oriented tasks yield better results than a meaning-oriented task. An important factor determining task effectiveness for vocabulary learning is the amount of word-related activity that the task induces. More words are acquired through tasks than through reading (Laufer, 2003). FonF (Focus on Form: drawing students' attention to linguistic elements as they arise incidentally in lessons), and FonFs (Focus on Forms: teaching discrete linguistic structures in separate lessons), whether related to a communicative task or not, play a crucial role in building the learner's lexical competence (Laufer, 2005).

Hill (2000) explored the usefulness of online tasks and concluded that contrary to Krashen's Input Hypothesis, the conventional comprehension task does not necessarily promote vocabulary learning. Tasks that require greater involvement with words result in better long term recall. The web is excelling as an interactive medium to present information in a more clearly comprehensible format. In addition, word-focused tasks (Laufer, 2003), form-focused instruction (Ellis, 2004), and form-oriented tasks (Hill & Laufer, 2003), work more effectively than natural acquisition through reading or meaning-focused instruction.

Since phonological skills and awareness are very important in early literacy, some studies on EFL children were conducted to measure how phonological ability affects second language vocabulary learning. Hu (2003) investigated the role of phonological memory and awareness for early childhood English learners in Taiwan. Phonological awareness is the ability to apprehend and manipulate smaller and smaller units of sound and facilitate the connection between letters and the sounds they represent in words. Phonological memory is the ability to hold sound-based information in immediate memory. Hu concluded that phonological memory is related to foreign language word learning, whereas phonological awareness is not. In contrast to this result, the study by Masoura and Gathercole (2005) showed that the children's speed of learning new English is independent of phonological memory skills. Their results show that the use of existing lexical representations is important as a means of supporting the acquisition of new vocabulary forms as well as increasing familiarity with the sound structure of a language.

In summary, the first trend of the EFL studies is examining whether Asian learners' characteristics are different from the other populations regarding the same instruction since many studies were conducted with Chinese students as English learners. The second major topic of the studies in this group is what kind of instruction supports vocabulary learning with the premise that vocabulary can be better acquired in task-related instruction rather than in incidental learning. Last, study results with child learners of EFL show that a meaning-related task is more related to vocabulary learning than phonological skill.

## **Review of Research on Vocabulary Instruction**

Petty, Herold, and Stoll (1968) analyzed 80 studies of vocabulary instruction in an attempt to identify the most effective methods of vocabulary instruction. They concluded that the data available at that time did not allow for a single “best “method of vocabulary instruction, but any instruction appeared better than no instruction. They also found that methods involving several techniques seemed more effective than those involving a single technique. In a later review, Mezynski (1983) found several factors related to the success of vocabulary instructional methods in improving reading comprehension. They were (1) the amount of practice given in learning the words, (2) the breadth of training in the use of the words, and (3) the degree to which active processing is encouraged.

Stahl and Fairbanks (1986) analyzed the effects of vocabulary instruction on the learning of word meanings and comprehension using a method of meta-analysis. They examined the effect of vocabulary pre-instruction as a way of improving reading comprehension and examined the components of effective vocabulary instruction. They concluded that the most effective vocabulary teaching methods included both definitional and contextual information in their programs, involved the students in deeper processing, and gave the students more than two exposures to the to-be-learned words. However, those three reviews of research (Petty, Herold, & Stoll, 1968; Mezynski, 1983; Stahl & Fairbanks, 1986) focused on first language vocabulary learning or instruction, not for ESL learners or second language learners.

The research related to second language vocabulary instruction has been devoted to discussing the relationship between individual learner factors and strategy use. In an

integrative review of 67 sample research reports on English as a second language learners' cognitive reading process, Fitzgerald (1995) suggested that experiences and prior knowledge affect comprehension and recall, and that vocabulary knowledge may be a highly significant variable in ESL learners' success. Fitzgerald also determined that more proficient ESL readers make better use of vocabulary knowledge than less proficient ESL readers. According to Green and Oxford (1995), learning style, motivation, belief, gender, career orientation, and personality also affect strategy use. While supporting these studies, Gao (2003) focused on the relationship between these relevant learner factors and changes in learners' strategy use of Chinese students who are studying in the UK.

Barcroft (2004) overviewed the major areas of research related to second language vocabulary acquisition, and summarized ten research areas. These include incidental vocabulary learning, lexical requirements for comprehension, input enhancement and text-based factors, vocabulary learning strategies, combined indirect and direct vocabulary instruction, methods of direct instruction, word-based determinants of learn ability, bilingual mental lexicon, receptive versus productive vocabulary knowledge, and lexical input processing. He also discussed five principles for effective second language vocabulary instruction with emphasis on lexical input processing. The instruction for second language learners should: (1) present new words frequently and repeatedly in the input; (2) use meaning-bearing comprehensible input when presenting new words; (3) limit forced output during the early stages of learning new words; (4) limit forced semantic elaboration during the initial stages of learning new words; and (5) progress from less demanding to more demanding vocabulary-related activities.

## Summary

In the first section of this chapter, the history of vocabulary instruction for English language learners was discussed in relation of literacy theories and practices from audio-lingual methods based on behavioral psychology to the communicative approaches. The debate between incidental learning and intentional learning was also discussed. Results of empirical study findings were presented. It was believed that regardless of the evident role of reading in vocabulary gain, the explicit vocabulary instruction could help learning vocabulary, especially for the English language learners.

The second part of this chapter described current characteristics of vocabulary instruction research. Which methods or strategies are beneficial in different settings have been examined using ESL or EFL learners by many researchers. There were some similarities and dissimilarities between the two groups of studies because their foci of research were varied, and there were some mixed results of the same methods of instruction. Finally, the reviews of the research that have most affected vocabulary instruction studies were discussed in order to cover some issues already reviewed by second language researchers.

This review provides a starting point to answer the research questions and resolving debates about vocabulary development for English learners. This also helps to build a theoretical background of vocabulary instruction as a basic knowledge to determine and categorize instruction.

## **CHAPTER III**

### **METHODOLOGY**

The review of the literature in the Chapter II yields inconclusive results with respect to the effectiveness of vocabulary instruction especially for English language learners. To better understand the general effectiveness of vocabulary instruction, combining empirical research is needed instead of another individual study. Therefore, a meta-analysis of the body of empirical research is best suited for my research purpose since meta-analysis is the statistical review of individual studies.

The characteristics of meta-analysis are discussed in the first section. First, backgrounds of meta-analysis are discussed to understand the history of this method. Strengths and weaknesses of meta-analysis are summarized as a quantitative review. The second section addresses the procedures of meta-analysis in five steps that I followed for the current study.

#### **Background of Meta-analysis**

Meta-analysis is a set of statistical techniques for combining information from different studies to derive an overall estimate of a treatment's effect. The modern era of meta-analysis began with the work of Glass (1976) in psychotherapy, since he defined meta-analysis as the statistical analysis of a large collection of analysis results from individual studies for the purpose of integrating the findings. However, the quantitative synthesis of findings from different studies can be found in 1904. Pearson invented a

statistical method for combining divergent findings by averaging the correlations between inoculation for typhoid fever and mortality for five separate samples (Cooper & Hedges, 1994).

Current meta-analytic techniques involve subtle and discriminating procedures, of which Pearson's averaging is one of the simplest (Hunt, 1997). In 1937 Cochran went off in a different direction and worked out a way of combining the sizes of the effects reported in studies rather than the correlations between treatment and effect; although this approach would become a key technique of meta-analysis, it also attracted little attention. But by the 1950s scientists increasingly needed to sum up numerous studies in their fields and reconcile their differences. Researchers began to develop methods to combine the results of studies within medicine, psychology, and sociology.

By the early 1970s, others were designing methods for aggregating studies of teaching methods, television instruction, and computer-assisted instruction. Rosenthal (as cited in Hunt, 1997) was developing a technique for combining the effect sizes of psychological studies. Glass (1976) was also working out a remarkably similar method of combining studies of the effects of psychotherapy without knowing Rosenthal's work. In April 1976, Glass chose to highlight a new and higher level of scientific analysis to which he gave the name "meta-analysis." From Glass's 1976 presentation, the prestige and practice of meta-analysis has grown steadily, with increasing speed and spreading from one discipline to another. In the 1980s, it was marked by the publication of a number of book-length expositions of the concepts, methods, and statistical theory of various versions of meta-analysis. Meta-analysis spread rapidly in the social sciences, especially education and psychology, and caught on with special vigor in the health sciences.



## **Literature Review of Meta-analysis**

Meta-analysis is one of the ways to review literature on scientific facts. Literature reviews play an important role in scientific development since science is a cumulative endeavor (Cook & Leviton, 1980). Both traditional literature review and meta-analysis require drawing up a list of theoretically relevant studies, and examining each study to determine a particular relationship related to research purposes. Many qualitative reviews share the same assumptions about the role of statistical significance testing with meta-analyses.

Meta-analysis tends to include many more studies than traditional reviews. Procedures of meta-analysis allow quantitative reviews and syntheses of the research literature, addressing some potential problems with traditional reviews. Using the traditional method of integrating research studies, a researcher provides a narrative, chronological discourse on previous findings. Wolf (1986) summarized the potential problems with traditional literature reviews into five points: (1) selective inclusion of studies, often based on the reviewer's own impressionistic view of the quality of the study, (2) differential subjective weighting of studies in the interpretation of a set of findings, (3) misleading interpretations of study findings, (4) failure to examine characteristics of the studies as potential explanations for disparate or consistent results across studies, and (5) failure to examine moderating variables in the relationship under examination.

In a meta-analysis, research studies are collected, coded, and interpreted using statistical methods similar to those used in primary data analysis. Meta-analysis allows

investigating the relationship between study features and study outcomes with coding the study features according to the objectives of the review and using statistical methods to show the relationships. This result is an integrate review of findings that is more objective and exact than a narrative review.

Green and Hall (1984) pointed out the advantages of meta-analysis as a way of literature review. First, this type of quantitative review is an efficient way to summarize large bodies of literatures. While the traditional reviewers cannot describe all the study designs and results, the quantitative meta-analysts can analyze 500 studies after they spend considerable time coding attributes of study outcome. Second, the meta-analytic review is likely to be more objective than the traditional review. The quantitative reviewer is more likely to find and include all relevant studies and is less likely to give undeserved weight to studies that are consistent with his or her prejudices. Third, statistical analyses can find relationships and trends too subtle to be seen otherwise. For example, the quantitative reviewer can investigate interaction and main effects which cannot be done with other types of reviews. The strength of the quantitative review is its ability to ask not simply whether there is overall support for a hypothesis, but also whether the extent of support varies with attributes of the studies. The ability to look for interactions in the data means that the quantitative reviewer can test hypotheses that were never tested in individual studies. Finally, the meta-analysis highlights gaps in the literature and is therefore a better source of insight into appropriate new directions for research.

## **Meta-analysis as a Quantitative Study**

Meta-analysis uses the quantitative methods to summarize and analyze research literature. It is called research integration (Walberg & Haertel, 1980) and quantitative assessment of research domains (Rosenthal, 1980). Green and Hall (1984) used the term ‘quantitative review’ to describe this kind of literature review. Both a quantitative study and a meta-analysis study use statistical methods to conduct a study. Statistical methods are useful in analyzing all kinds of data in the social and behavioral sciences. Like a quantitative study, the major problems in meta-analysis are choosing the appropriate dependent and independent variables and analyzing data with the needed statistical analysis.

Meta-analysis can be very useful when used wisely like any statistical analysis (Cook & Leviton, 1980). Many of the standard methods of statistical analysis are useful in meta-analysis. Most quantitative reviews use some index of the size of the effects, which not only helps to establish the statistical significance of the results, but also describes their magnitude. Meta-analysis can be understood as a form of survey research in which research reports, rather than people, are surveyed (Lipsey & Wilson, 2001). A coding form (survey protocol) is developed, a sample or population of research reports is gathered, and each research is “interviewed’ by a coder who reads it carefully and codes the appropriate information about its characteristics and quantitative findings. The resulting data are then analyzed using special adaptations of conventional statistical techniques to investigate and describe the pattern of findings in the selected set of studies.

Since meta-analysis is only one of many ways to summarize, integrate, and interpret selected sets of scholarly works, there needs to be a specific condition under which meta-analysis can be conducted. First, meta-analysis applies only to empirical research studies; it cannot be used to summarize theoretical papers, conventional research reviews, policy proposals, and the like. Second, it applies only to research studies that produce quantitative findings, that is, studies using quantitative measurement of variables and reporting descriptive or inferential statistics to summarize the resulting data. Third, meta-analysis is a technique for encoding and analyzing the statistics that summarize research findings as they are typically presented in research reports. If the full data sets for the studies of interest are available, it will generally be more appropriate and informative to analyze them directly using conventional procedures rather than meta-analyze summary statistics (Lipsey & Wilson, 2001).

### **Advantage of Meta-analysis**

To answer my research questions, I chose standardized mean differences as a statistical measure in my meta-analysis. Most of the studies I collected for meta-analysis have means and standard deviations for both control groups and experimental groups. Using this meta-analysis technique, the research relating to vocabulary instruction for English language learners was synthesized. In this synthesis, many experimental and quasi-experimental studies which were excluded by the previous meta-analyses of vocabulary research as well as the NPR report of vocabulary instruction (2000) were reviewed. The variety of research methods, different samples (i.e., from children to

adults, from beginners to advanced learners, and from living in English speaking country to the other countries which speak other first languages) and size (i.e., number of sample size vary from around 20 to 200), criterion measures, and statistics affected the experimental study results. It could be assumed that different instruction and setting factors would have different effects on acquiring vocabulary for English language learners. Nonetheless, with these kinds of differences, by employing the meta-analysis techniques, it was possible to explore the effects of vocabulary instruction on the relation of learners of different ages, proficiency levels, language groups, and social setting.

### **Procedure**

The meta-analytic process has five basic phases (Glass, 1976). First, formulating the problem; deciding what questions the meta-analyst hopes to answer and what kinds of evidence to examine. Second, collecting the data; searching for all studies on the problem by every feasible means. Third, evaluating the data; deciding which of the gathered evidence is valid and usable; eliminating studies that do not meet these standards. Fourth, synthesizing the data; using statistical methods, such as the combining of probabilities and the combining of effect sizes, to reconcile and aggregate disparate studies. Finally, presenting the findings; reporting the resulting “analysis of analyses” to the wider research community, providing details, data, and methods used. I will conduct my study based on these steps.

**1. The literature search.** A search using an online database through the Texas Tech Library was initiated. Three major databases were Dissertation Abstract, Linguistics and Language Behavior Abstract (Cambridge scientific abstract), and ERIC First Search. Using the term “vocabulary instruction,” “vocabulary learning,” “English language learners,” “English language as a second language,” and “English language as a foreign language” yielded many citations, but since meta-analysis needs experimental or quasi-experimental studies, citations without quantitative methodology were removed. After read more than 500 abstracts related to those search, I could print about 100 articles from electronic journals on the above databases or offline journals from the Texas Tech Library, and about 40 articles through the document delivery system. I also borrowed 17 dissertations from interlibrary loans and purchased 5 dissertations.

**2. Selection of the studies.** This study is a quantitative analysis of a sample of studies that address the effects of vocabulary instruction on English language learners. That sample of studies includes articles published in journals and dissertations. These studies report the results of observations of vocabulary instructions after 1975. In order to be included in this meta-analysis, the criteria for selection of the studies were as follows: (1) studies must have involved students learning English vocabulary, (2) studies must have reported on measured outcomes for groups of students with “special” instructions for learning or groups of students being taught “innovative” teaching methods which are normally called an experimental group and the studies must have reported on groups of students in comparison instruction or in comparison groups receiving traditional instruction as a control group , (3) when a single article gave results for several results

like pretest, post test, only the result for the post test were included, and if an immediate test and delayed test were conducted, the average scores of both tests were calculated, and (4) the study must not have made obvious reporting errors.

**3. Coding of the studies.** Coding studies for meta-analysis is essentially like doing a survey. In survey research, it is important to prepare the questionnaire carefully, train the interviewers (coders), and monitor the completeness, reliability, and validity of the data that are obtained (Lipsey & Wilson, 2001). Two parts of coding protocol, the part with information about study characteristics (study descriptor) and the part with information about the empirical finding of the study (effect size) should be included in meta-analysis. Study findings are the dependent variables of meta-analysis while study characteristics such as methods, measures, samples, constructs, treatment, context are the independent variables of a meta-analysis.

Information coded for each effect included the author's last name, publication year, publication types, dependent measures, basic methodology, and sample characteristics. Sample characteristic include age (children, elementary students, secondary students, and college students), geographical location (the United State, Canada, Europe, Asia, and other), and language level (beginners, intermediates, and advanced learners). Dependent measure refers to the types of instructional approaches, which are contextualizing, semi contextualizing, decontextualizing, using multimedia, and using first language support. Table 3.1 shows the categories for my coding information. Coding sheet was attached in Appendix A.

Table 3.1 Coding Categories Used In Analysis

Category	Subcategory
Outcome	achievement score
Publication type	dissertation journal articles
Publication date	
Region where research took place	ESL vs. EFL  The United States Canada Europe Asia Other
Participant age	children Elementary students Secondary students College students
Language level	beginners intermediates advanced learners not mentioned
Instructional approach	contextualizing instruction semi-contextualizing instruction decontextualizing instruction  using multimedia :mediating computer, audio, video using first language support :translation and dictionary
Vote count	significant negative non-significant negative non-significant positive significant positive



As I explained in chapter I, based on previous literature by Oxford and Crookall (1990) as well as the National Reading Panel(2000), the collected data in this study was divided into three categories :(1) contextualizing (incidental acquisition from reading), (2) semi contextualizing (semantic elaboration),and (3) decontextualizing (focus on form and word definition). In addition to those three categories of methods, the use of multimedia (mediating computer, audio, video) were considered as one of instructional methods as well as the use of first language support (translation, dictionary). The multimedia methods -mostly with computers- or programs are used with one of three instructional approaches. For example, one computer vocabulary learning program used mainly decontextualizing methods, the others are for semi-contextualizing or contextualizing. Therefore, I coded one study with three instructional levels, first which instructional approaches are used among three (decontextualizing, semi-contextualizing, and contextualizing) second, whether the multimedia is used, and whether first language supports are exist.

When conducting a meta-analysis, it is necessary to consider the reliability of locating research results. Coder reliability has two dimensions: the consistency of a single coder from occasion to occasion, and the consistency between different coders (Lipsey & Wilson, 1995). In my study, to retain the consistency of coding from the beginning to the end, I coded again the studies I had coded earlier and after sufficient time had passed without reference to the original coding.

To generate a relatively stable reliability estimate, Lipsey and Wilson (2001) suggest having 20 or more study samples with 50 + being more desirable. For a small meta-analysis, it may be necessary to use all the studies in the reliability check. Since my

total number of studies was little more than 40 studies, I did all the studies in the reliability check. Several changes were made when I found mistakes. I compared the results before and after some changes.

After I checked with all of my samples for reliability check, the other coder drew ten percent of the total samples and coded for reliability check. Generally, coder reliability is checked by drawing a sub sample of the coded studies, having the other coder code them again, and comparing the results. For this kind of reliability check for my study, one of my dissertation committee members helped me to code some sample of studies. He coded 10% of my sample studies, and calculated the inter-rater reliability. The score of inter-rater reliability was .97, it could be concluded that the coding process is reliable. Generally if the score of inter-reliability is the above .80, it is reliable statistically.

Another reliability issue is the degree of consistency in calculating and recording the significance levels and effects size estimates of studies. I examined the homogeneity of effect sizes; homogeneity indicates whether the distribution of the effect sizes accurately reflects the mean effect size for the population (Lipsey & Wilson, 2001). The Q statistic is computed in a test of homogeneity and is distributed as a chi-square. When tests of homogeneity are significant, it is often a result of methodological differences or participant differences, which means that different factors are associated with these effect sizes. In addition, I considered whether outliers should be noted and possibly removed from the analyses if there are some outliers which are beyond two or three standard deviations of the mean. When data exploration was complete, weights were calculated since bias could result from studies with smaller sample sizes. Weighted values could be

calculated by using the inverse variance weight formula to correct bias (Hedge & Olkin, 1985; Lipsey & Wilson, 2001).

**4. Measuring study findings.** My entire sample articles are experimental studies comparing achievement scores among different vocabulary language learning conditions. Most of the studies compare which instruction works more effectively than others with two or three comparisons. Three comparison studies are divided into two studies; two pairs of one control group and one experimental group. Therefore, there is more than one study from one article in this meta-analysis. Mostly I coded a traditional /old method as a control group and a new method, which researchers usually wanted to measure the effect, as an experimental group. Additionally, I used posttest scores if the studies reported pretest and posttest scores.

Research findings for achievement scores are mostly the vocabulary test scores after experiments. The finding in each study must be encoded a value on the same effect size statistic. Research findings in the form of group contrasts involve a variable that is measured on two or more groups of respondents and then compared across groups. The descriptive statistics that typically characterize this situation are the means, standard deviations, and sample size for each group on each variable. Differences among groups of these kinds are often examined with such familiar statistical tests as the t-test, analysis of variance (ANOVA), chi-square, and the like (Lipsey & Wilson, 2001). Most experimental and quasi-experimental studies used in this analysis provide findings in this form with a comparison of the experimental group with the control group on one or more dependent variables.

Independent variables are all the study characteristics which can be coded for each study and their effects determined by statistical analysis. Green and Hall (1984) mentioned that meta-analysis should do more than establish the overall existence and size of effect. Many aspects of the studies can be examined if there are enough studies. The most frequent criticism of quantitative review is that findings cannot be combined uncritically. However, the quality of the study, the type of analysis done, the nature of the sample, even the age of the study can be weighed in a meta-analysis. All of these characteristics can be treated as independent variables, which can be coded for each study and their effects determined by statistical analysis rather than by subjective impression.

Lipsey and Wilson (2001) suggested a set of categories that identify study descriptors of different types, which are independent variables in other words. The first category is the substantive aspects of the study, i.e., the nature of the persons in the samples, treatments applied, cultural or organizational context, and so forth. These independent variables are most interesting when the meta-analyst attempts to account for different results across studies. My independent variables in this category are the age of samples and their language levels as well as the location of study conducted.

The second category of independent variables is study methods and procedures. Wilson (1995) states that study findings are often related to methodological differences among the studies. If methodological differences are not coded and examined, the meta-analyst might easily misinterpret their effects as a substantive difference. Therefore, the statistical power and the types of statistics are in this category. In addition, the most important independent variables in this study are different kinds of vocabulary instruction

as methods of assignment to examine whether some instruction produces larger effects than others.

The final category of independent variables is called source descriptors by Lipsey and Wilson (2001). This category includes factors that have to do with the general study context, the particulars of the publication, researchers, and so forth that are neither substantive nor methodological. My independent variables in this category are the year of publication, publication form, and names of researchers.

I could question whether the age of samples, their levels of English proficiency, and geographical location, mediated the relation between vocabulary instruction and achievement scores. Average effect sizes across studies could be calculated and evaluated for these potential mediating variable categories. It is important to test whether these third set of variables are mediators of the effect in my primary research hypothesis. For examples, suppose that the average effect size for ESL condition is .50 ( $SD=.09$ ,  $n=15$ ), and for EFL condition is .80 ( $SD= .08$ ,  $n=17$ ). The effect size for each of these 43 studies could be correlated with the geographical location of study conducted. This result indicates that stronger effect sizes tended to be obtained from EFL samples rather than ESL samples. That is, the relation between vocabulary instruction and achievement scores tended to be stronger for EFL samples than for ESL samples.

In addition, the publication source of study could mediate the effect sizes. Kulik and Kulik (1989) considered the effects of publication source as a factor related to study outcome. Results found in journal articles are usually more positive than are result from dissertations. Since study effects are somewhat higher in journal articles than in dissertations, it could be concluded with higher effect sizes than actually exists if I only

use journal articles. Therefore, I did not limit my selection of study to one type of source. I tried to collect not only journal articles but also dissertations.

**5. Analysis.** Most of the studies I collected for meta-analysis have means and standard deviations for both control groups and experimental groups. To summarize statistically the results of the various studies, it was necessary to express these results in terms of a common scale, which is effect size in a meta-analysis. To calculate the effect size in this study, the difference between two group means was divided by the pooled standard deviation (Wolf, 1986), noted in Formula 1. Effect size is

$$d = \frac{\bar{X}_t - \bar{X}_c}{S_p} , \quad (1)$$

where  $\bar{X}_t$  is the mean for a treatment or experimental group,  $\bar{X}_c$  is the mean for a control group, and  $S_p$  is the pooled standard deviation (Hedge & Olkin, 1985). Cohen (1988) defined  $d$  as the difference between the means  $M_1 - M_2$ , divided by standard deviation of either group. He stated that the standard deviation of either group could be used when the variances of two groups are homogeneous. In meta-analysis the two groups are considered to be the experimental and control groups, so that the pooled standard deviation is commonly used. The pooled standard deviation is found as the root mean square of the two standard deviations.

After calculating the effect size  $d$  for each of the independent studies, the average of these effect sizes was calculated using Formula 2. This represents the estimate of effect size across all the studies,

$$d_{average} = \frac{\sum d}{n} \quad (2)$$

where  $d$ =the effect size for each independent study,  $n$ =the number of studies.

Positive values of effect size are associated with results favoring the experimental group, while negative values are associated with results favoring the control group.

Because the sample sizes differ from study to study, the weighted effect sizes need to be calculated to give more precise estimate to large studies. This was calculated by formula 3, weighting the more precise estimate when pooling (Hedge & Olkin, 1985).

$$w_i = \frac{\tilde{n}_i}{\sum_{j=1}^k \tilde{n}_j}$$

$$\tilde{n} = \frac{n_t n_c}{n_t + n_c} \quad (3)$$

Where the pooled weighted estimator is

$$d_w = w_1 d_1 + \dots + w_k d_k.$$

After calculating the overall effect size of the total study, the effect sizes of sub-group are calculated. To know the all studies and the sub-groups shared a common effect size for the dependent variables, the homogeneity statistics are calculated (Hedges & Olkin, 1985). The formula 4 for the homogeneity statistic (Q-total) is (Berley, 1993)

$$Q_T = \sum_{i=1}^k d^2 \bar{\partial}_d^2 - \frac{\left( \sum_{i=1}^k \frac{d}{\bar{\partial}_d^2} \right)^2}{\sum_{i=1}^k \frac{1}{\bar{\partial}_d^2}} \quad (4)$$

Using the Chi-square table, the homogeneity of effect sizes is determined. If Q-total is not larger or is statistically non-significant, the single effect size fits the data for meta-

analysis. A significant value of chi-square indicates that the effect sizes are not homogenous. Appendix B shows the SPSS syntax used in this study.

### ***Pilot Study***

Based on preliminary data analysis of this study, I could state that the effect sizes for all studies are positive values, which means all of these experimental groups are more effective than the control groups on vocabulary learning. In addition, based on the preliminary finding for  $d$  (average)=.94, this meta-analysis could conclude that vocabulary instruction related to any instructional approaches is more effective than vocabulary instruction related to traditional instruction or no instruction by approximately .94 standard deviation units. This would be the best estimate regardless of how individual studies measured their variables in many studies.

Cohen (1977) provides rough guidelines of  $d = .2$  (small effect),  $d = .5$  (medium effect), and  $d = .8$  (large effect). Therefore, the  $d = .94$  in this study were considered to be significant. That is, vocabulary instruction would be effective in practical ways since a 0.50 standard deviation improvement in achievement scores is considered to be a conventional measure of practical significance (Rossi and Wright, 1977).

Inferential statistics are used to test the hypothesis and make inferences from the sample about the population from which it was drawn. Glass, McGaw, and Smith (1981) claimed that the typical meta-analysis virtually never meets the condition of probabilistic sampling of a population. They took the position that inferential statistics has little role to play in meta-analysis. Under a random sample of some hypothetical population, inferential techniques are applied and the results inspected. In all other instances, it will



be up to the meta-analyst to decide whether to apply them and how to interpret them (Rudner, Glass, Evatt, & Emery, 2000). In other words, the inferential statistical techniques in meta-analysis are easy to misapply and misuse. However, conducting a mediating effect analysis provides some evidence to interpret meta-analysis results appropriately. I used Hedge's homogeneity tests which address inferential issues as one of the inferential techniques to overcome the limitation of the samples in this meta-analysis.

My preliminary study result showed that this study was not homogeneous. Heterogeneity is a warning that it may not be appropriate to combine and synthesize all the results in one meta-analysis (Wolf, 1986). In this case, it is better to conduct mediating effects analyses until the results shows homogeneity. Therefore, one instructional approach has been selected to find homogeneity result. However, there was also heterogeneity in this mediating analysis. Opposed to Hedges (1982) and Hunter Hunter, Schmidt, and Jackson (1982) suggestion that it is inappropriate to include all heterogeneity studies in one meta-analysis, Harris and Rosental (1986) argue that heterogeneity is common whenever many studies by different investigators using different methods are examined. This meta-analysis investigates different vocabulary instruction for English language learners, thus, it is likely to be heterogeneous due to the different sample characteristics and the different variable characteristics. It is also possible to explain that this kind of heterogeneity is common because there is no agreement with one best method among all kinds of different instructional approaches.

## **Summary**

This methodology provided systematic procedures for answering my research questions. First, how can the vocabulary instruction for English learners be categorized? Using a detailed coding process, the vocabulary instruction studies were grouped into some distinct characteristics. This revealed the wide range of vocabulary instruction most commonly used in the classrooms and researched by second language professionals.

Second, to what degree is this instruction successful? Effect size was the prime estimate of degree. Effect size estimates were computed on reported achievement.

Third, the mediating effects of this study were found using a homogeneity test. The effects of each group of categories were compared in order to determine the effectiveness of each vocabulary instruction.

## **CHAPTER IV**

### **PRESENTATION OF RESULTS**

This study was designed to examine a wide selection of vocabulary instruction for English language learners in order to determine effective instruction based on learner characteristics. This chapter is divided into two major sections. In the first section, an overall broad picture of the studies that investigated the effectiveness of vocabulary instruction is presented. This first section also presents the research synthesis. The second section presents the statistical results to categorize these studies, to determine the effect sizes they produce, and to note any factors that might have affected effectiveness of vocabulary instruction or study outcomes. This second section directly addresses the research questions of this investigation. This analysis provides a general landscape of characteristics of vocabulary instruction and the effectiveness of that instruction.

#### **Description of the Studies**

In this section, different study features are synthesized and compared in order to provide some of the fundamental characteristics of the studies in the present meta-analysis. Thus, features such as research publication, research setting, and learner characteristics, and research design are discussed here.

## **1. Research Publication**

Forty-three studies published between 1985 and 2006 (except one in 1975) were included in the current meta-analysis. Those studies met the inclusion criteria, were published in 30 different articles and dissertations (See Appendix C). Some articles or dissertations provided more than one experiment, so those articles and dissertations make up the additional 13 studies.

Although the database search was done without specifying dates, there was only one study from the 1970s and the other one from the 1980s that met the inclusion criteria. Except for two reports, most of them were published after 1990, especially 17 articles and five dissertations were published after 2000. Of the 30 publications, seven were dissertations while the other 23 were journal articles, published in 16 different journals. I found four articles in *Language Learning*, three in *The Canadian Modern Language Review*, two in *Studies in Second Language Acquisition*, two in *Computer Assisted Language Learning*, one in each of these journals, *IRAL( International Review of Applied Linguistics, Language , Learning & Technology, , EUROSLA Yearbook, Journal of Experimental Education, TESOL Quarterly, Asian Journal of English Language Teaching, English for Specific Purposes, CAELL Journal :Computer-assisted English Language Learning, Applied Linguistics, Review of Applied Linguistics, CAUCE ( Spain), Computer in Human Behavior.*

## **2. Research Setting**

The features of the research settings included whether the study was conducted with college, secondary, or elementary students, in EFL or ESL setting. Of the 43 studies,

36 (or 84%) were conducted in an EFL setting, seven (16%) in an ESL setting. In detail, the six studies (14%) in an ESL setting were conducted in the US except one in New Zealand. The 25 studies (58%) out of 36 in an EFL setting were conducted in Asia while the six studies (14 %) in Europe. Twenty three studies (53%) were conducted with university or college students, 11(26%) with secondary school students, and five (12%) in an elementary setting, and three (7%) with early childhood children.

Fourteen studies (33 %) among 43 used multimedia supports, mostly computer programs. Twenty five studies (58%) conducted in EFL conditions allowed first language translations or bilingual dictionary look-up during experiments.

### ***3. Learner Characteristics***

Learner characteristics are important factors to synthesize. A total number of 2991 subjects participated in this analysis with various first languages and various proficiency levels. Table 4.1 summarizes their first languages, levels of proficiency, and their ages. Most studies conducted in the US were summarized with mixed languages in this table because ESL students in the US were from different language backgrounds. However, the majority student population of this group came from various Asian countries even though there were some students that came from the other parts of the world. The fact that most studies involved Asian students might be related to the characteristics of ESL population of college students in the US. Except those studies conducted with ESL, most of the studies have a homogeneous first language group. Since many studies were conducted in Saudi Arabia and Israel as well as China, Arabic, Hebrew, and Chinese are the major first three languages in this analysis.

**Table 4.1 Learners Characteristics**

<i>variable</i>	<i>Number of studies</i>	<i>details</i>
First Language	6	Arabic
	9	Chinese
	3	Dutch
	4	Hebrew
	1	Japanese
	2	Korean
	2	Russian
	2	Spanish
	1	Thai
	1	Turkish
	7	Mixed languages in ESL
English proficiency level	3	Beginner
	9	Intermediate
	6	Advanced
	22	Not reported
Participant age	3	Children
	5	Elementary students
	11	Secondary students
	25	University & above

Most studies conducted in the ESL condition reported the level of English proficiency because the student language level in ESL classes would easily be known. However, it might not be easy to report the student level of English proficiency in EFL condition especially with elementary or secondary students. Instead of reporting the level of English proficiency, researchers replaced that with the duration of English learning. Therefore, I could not decide their language level on my own so I left them in the “not reported” in Table 4.1.

In addition, more studies were conducted with university students compared to the other population of students. One of the reasons might be that researchers could more easily conduct their experiments with college students as samples. The order the participants, the more studies were conducted.

#### ***4. Research Design***

All the studies reported the sample size (N size). The study with the smallest number of participants had 29, while the one with the largest had 293 participants. The average N across all the 43 studies was 70. Many studies reported the scores of pretest and posttest, but only posttest scores were used for the analysis. If the immediate test and delayed test were conducted, the average scores of both tests were calculated. The studies used various types of test formats. The formats were mostly multiple choice tests, cloze questions or mixed formats of multiple choice items and cloze test, as well as writing a sentence or essay.

The collected studies used different types of treatments. There are three categories of different instruction that were used in the learning of new words: decontextualizing, semi-contextualizing and contextualizing. Studies differed in their choice of control or comparison treatment as well. Nine studies had a true control group, 33 made a comparison between various instructions.

## **Meta Analysis of Study Findings**

### ***1. Overall Effect of Vocabulary Instruction***

The overall effect size mean for the 43 studies of vocabulary instruction was  $d=.69$  with a total sample size of 2991. Findings displayed in Table 4.2 show that there is an overall large effect of vocabulary instruction for English language learners on the tests. In other words, the treatment groups performed .69 standard deviation units better than the control groups. This would be the best estimate regardless of how individual studies measured their variables in 43 studies.

Cohen (1977) provides rough guidelines of  $d= .2$  (small effect),  $d= .5$ (medium effect), and  $d= .8$  (large effect). Effect sizes can also be thought of as the average percentile standing of the average treated (or experimental) participant relative to the average untreated ( or control) participant. For example, an effect size 0 indicates that the mean of the experimental group is at the 50<sup>th</sup> percentile of the untreated group. An effect size of .8 indicates that the mean of the experimental group is at the 79<sup>th</sup> percentile of the



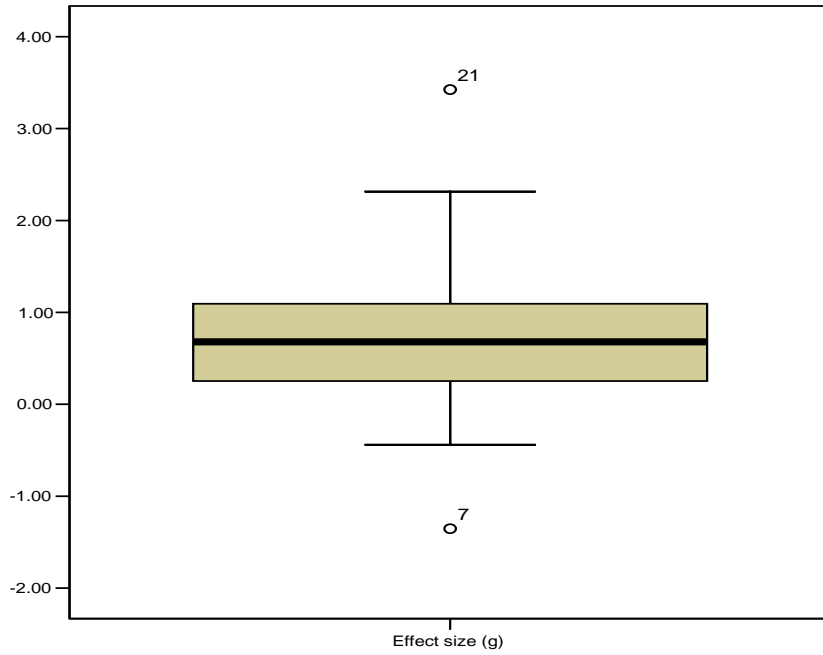
**Table 4.2. Overall Effect Size Mean for the 43 Studies of Vocabulary Instruction**

<i>Study</i>	<i>Weight</i>	<i>d</i>	<i>SE</i>	<i>95% Lower</i>	<i>95% Upper</i>
1	1.21	.67	.35	-.2	1.37
2	1.54	1.34	.31	.73	1.96
3	1.81	1.93	.29	1.36	2.49
4	1.32	2.31	.34	1.65	2.98
5	2.37	.55	.25	.06	1.05
6	1.31	-.27	.34	-.93	.40
7	.99	-1.35	.39	-2.12	-.59
8	1.97	1.07	.28	.53	1.62
9	1.24	-.25	.35	-.93	.44
10	.68	2.25	.47	1.33	3.18
11	2.82	.59	.23	.14	1.04
12	2.21	.77	.26	.26	1.29
13	1.15	.69	.36	-.03	1.40
14	1.53	1.35	.31	.73	1.97
15	3.12	.55	.22	.12	.98
16	5.17	1.11	.17	.78	1.45
17	2.38	-.09	.25	-.58	.41
18	2.33	.26	.26	-.24	.76
19	2.16	-.44	.27	-.96	.08
20	2.83	.37	.23	-.09	.82
21	.82	3.42	.43	2.58	4.26
22	3.53	.42	.21	.01	.82
23	2.30	1.71	.26	1.21	2.21
24	5.30	1.00	.17	.67	1.33
25	1.53	-.40	.32	-1.02	.21
26	1.85	.20	.29	-.37	.76
27	1.69	.95	.30	.37	1.54
28	1.10	.48	.37	-.25	1.20
29	10.58	.71	.12	.47	.94
30	.82	1.84	.43	.99	2.68
31	6.78	.47	.15	.18	.77
32	1.49	-.20	.32	-.82	.42
33	.96	.99	.40	.21	1.77
34	1.79	.65	.29	.08	1.22
35	1.60	1.25	.31	.64	1.85
36	3.63	.08	.20	-.32	.48
37	3.33	.84	.21	.42	1.26
38	1.80	1.97	.29	1.40	2.54
39	1.83	.71	.29	.15	1.28
40	1.92	.45	.28	-.10	1.00
41	1.79	.25	.29	-.32	.82
42	1.74	.06	.30	-.52	.63
43	1.67	.77	.30	.18	1.36
<b>Total</b>	100.00	.69	.04	.61	.76

untreated group. An effect size of 1.7 indicates that the mean of the experimental group is at the 95.5 percentile of the control group. Therefore, the effect size  $d = .69$  ( $SE = .04$ ) in this study indicates that the mean of the treatment group is at around the 76<sup>th</sup> percentile of the comparison group.

In addition to the effect size, the computation of a confidence interval helps to interpret a meta-analysis. According to Lipsey and Wilson (2001), confidence intervals indicate the range within which the population mean is likely to be, given the observed data. The confidence interval is useful in indicating the degree of precision of the estimate of the mean effect size. If the confidence interval does not include zero, then the mean effect size is statistically significant at the level specified by the confidence interval (Lipsey & Wilson, 2001). To construct the confidence interval, the standard error (SE) needs to be multiplied by a critical  $z$ -value representing the desired confidence level and add the product to the mean effect size for the upper limit, and subtract the product from the mean effect size for the lower limit.

For this study, the 95% confident interval ranged from the lower CI  $d = .61$  to the upper CI  $d = .76$ . Chi-square of heterogeneity statistic was 279.90, which means this meta-analysis is not homogeneous. Heterogeneity is a warning that it may not be appropriate to combine and synthesize all the results in one meta-analysis (Wolf, 1986). In this case, it is better to conduct mediating effects analyses until the results shows homogeneity. Therefore, the meta-analysis without outliers has been analyzed, since the graph displays two extreme cases in figure 4.1. These two studies should be considered outliers because these are above two standard deviations which is normally considered as outliers.



**Figure 4.1. Display of effect sizes from studies of vocabulary instruction**

Figure 4.1 also shows the mean effect size with confidence intervals as well as the number of studies of outliers. Therefore, the study ID 7 and 21 were deleted from further analysis. Therefore, the final number of studies included in this meta-analysis was 41. The mean effect size for the 41 studies of vocabulary instruction without outliers was .69 with a total sample size of 2902.

However, there was still heterogeneity in the meta-analysis of 41 studies: chi-square of heterogeneity statistic was 212.42, standardized error was .04. The heterogeneity was not only from the outliers.

In general, the publication source of study could mediate the effect sizes. In the current study, the mean effect size for 29 studies from journals was  $d=.80$  ( $SE=.05$ ) while the mean effect size for 12 studies from dissertations was  $d=.40$  ( $SE=.07$ ). Kulik and Kulik (1989) considered the effects of publication sources as a factor related to study outcomes. Results found in journal articles are usually more positive than are result from dissertations. Since study effects are somewhat higher in journal articles than in dissertations, effect sizes could be inflated. In fact, one of the problems with meta-analysis is that researchers can only analyze the studies that have been published. It is impossible to know how many studies that did not find significant effects have not been published. That is called the file drawer problem. If those studies had been published then the effect sizes would be smaller.

To avoid a bigger effect size than real for the further study, the study ID 7 and 21 were removed, so the final number of studies used in this meta-analysis was 41 and the overall effect size of this meta-analysis was .69 even though this effect size might be bigger than the real because of publication bias and the file drawer problem.

The coding procedure allowed me to compare vocabulary instruction studies by categories such as instructional approach, the place of the study, participant age, language levels, the effect of using multimedia, and the effect of first language support as well as publication types. Therefore, a comparison was made between groups to determine the effect size of each group and find the factors that affected the effectiveness of vocabulary instruction (I attached the effect sizes of each groups with the detailed information in Appendix D).

## ***2. Effects of Instructional Approaches***

Oxford and Crookall (1990) evaluated more than a dozen vocabulary learning techniques for students of second and foreign languages. These techniques are divided into four categories: decontextualizing, semi-contextualizing, fully contextualizing, and adaptable techniques. It was found that all instruction across 41 studies could be classified into three categories, decontextualizing, semi-contextualizing, and fully contextualizing because adaptable techniques reinforce those three categories of instruction.

According to Oxford and Crookall (1990), decontextualizing instruction removes the word as completely as possible from any communicative context that might help the learner remember. The example of this instruction is using word lists or flash cards where learners do not need much context to learn vocabulary. Rote memorization is an important part of learning new words and looking up words in dictionaries can help learners remember the meanings. The studies conducted in this groups measured the

effectiveness of flashcards (Sun & Dong, 2004) and word lists (Laufer, 2005), or measured the instruction focusing on word definition (Yeung, 1999), core meaning (Vespor & Lowie, 2003), or some other types of instruction related to word-focused tasks or form-focused tasks (Hill & Laufer, 2003). The effectiveness of using a dictionary was also measured (Oskarsson, 1975; Lopez, 2001).

Twelve studies were in this group with a total sample size of 1303. The effect size was  $d=.76$  ( $SE=.06$ ), which is larger than the overall effect size of the total studies. In other words, the effect of decontextualizing instruction is better than the other instruction. With a small variability ( $SE=.06$ ), the effect ranges from a lower CI  $d= .65$  to the upper CI  $d=.87$ , and chi-square statistic is 49.81. All the studies except one in this category were conducted in EFL in Asia and Europe.

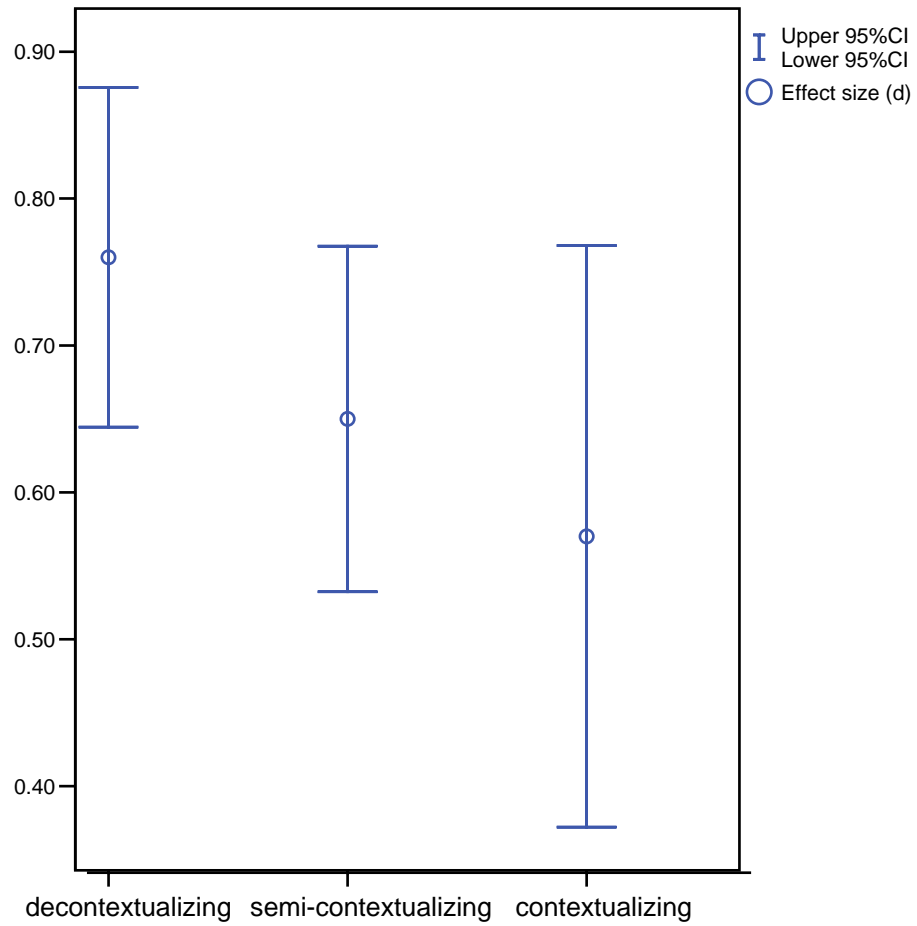
The second category of instructional approach is semi-contextualizing which allows some degree of context but falls short of full contextuality; thus new words may be linked with something that is meaningful to the learner, but they are not used as part of naturalistic communication. Word groupings, word or concept association, visual imagery, aural imagery, keyword, physical response, physical sensation, and semantic mapping are the examples of semi-contextualizing techniques by Oxford and Crookall (1990).

Some studies in this group measured the effect of keyword methods (Bird & Jacobs, 1999; Bussel, 1994; Faeqehi, 2003), the effect of semantic mapping (Srinaowaratt, 1991; Crow & Quigley, 1985), the effect of video or pictures (Al-Seghayer, 2001; Kim, 2006; Liu, 2006; Kang, 1994), and the other related instruction of semi-contextualizing. The mean effect size of 20 studies in this group was  $d=.65$  ( $SE=$

.06) with a sample size of 1145. The range of confidence interval is from a lower  $d=.53$  to the upper  $d=.78$  with the chi-square 96.09.

The final category of instructional approach is the contextualizing instruction. These fully contextualizing techniques embed the new words in a normal communicative context. Since a basic assumption of this approach is that massive amounts of reading would automatically increase their vocabulary, reading and listening or speaking and writing practice are emphasized. Nine studies were conducted to measure the effectiveness of contextualizing practice. The effects of composition or writing (Hulstijn & Laufer, 2001; Laufer, 2003), the effects of incidental context learning (Rodriguez & Sadoski, 2000; Sun & Dong, 2004), the effects of reading and retelling (Joe, 1996) and the effect of text writing compared to sentence writing (Al-Hadlaq, 2003) were measured. The effect size of this group was  $d=.57$  ( $SE=.10$ ) with a sample size of 454 in nine studies. Findings show that the contextualizing instruction yields the smallest effect size among three instructional groups with a larger variability ( $SD=.10$ ). The margin of error ranges from a lower  $d=.38$  to  $d=.77$  with a chi-square 63. Relatively low numbers of studies included in this group might be responsible for this wide confidence interval and does not allow trustworthy conclusions.

Figure 4.2. displays group comparisons of effect of different types of vocabulary instruction. Even though there seems to be a difference among three groups from the lowest effect size of contextualizing group  $d=.57$  to the highest effect size of decontextualizing group  $d=.76$  in Figure 4.2, the confidence intervals of the three groups overlap each other. Therefore, it can be concluded that the three types of different instruction did not yield significantly different results at this point.



**Instructional Approach**

	<i>d</i>	<i>SE</i>	<i>k</i>	<i>n</i>	<i>95 % lower</i>	<i>95% upper</i>
decontextualizing	.76	.06	12	1303.	.64	.88
Semi-contextualizing	.65	.06	20	1145	.53	.77
contextualizing	.57	.10	9	454	.37	.77

**Figure 4.2. Effects of instructional approaches**

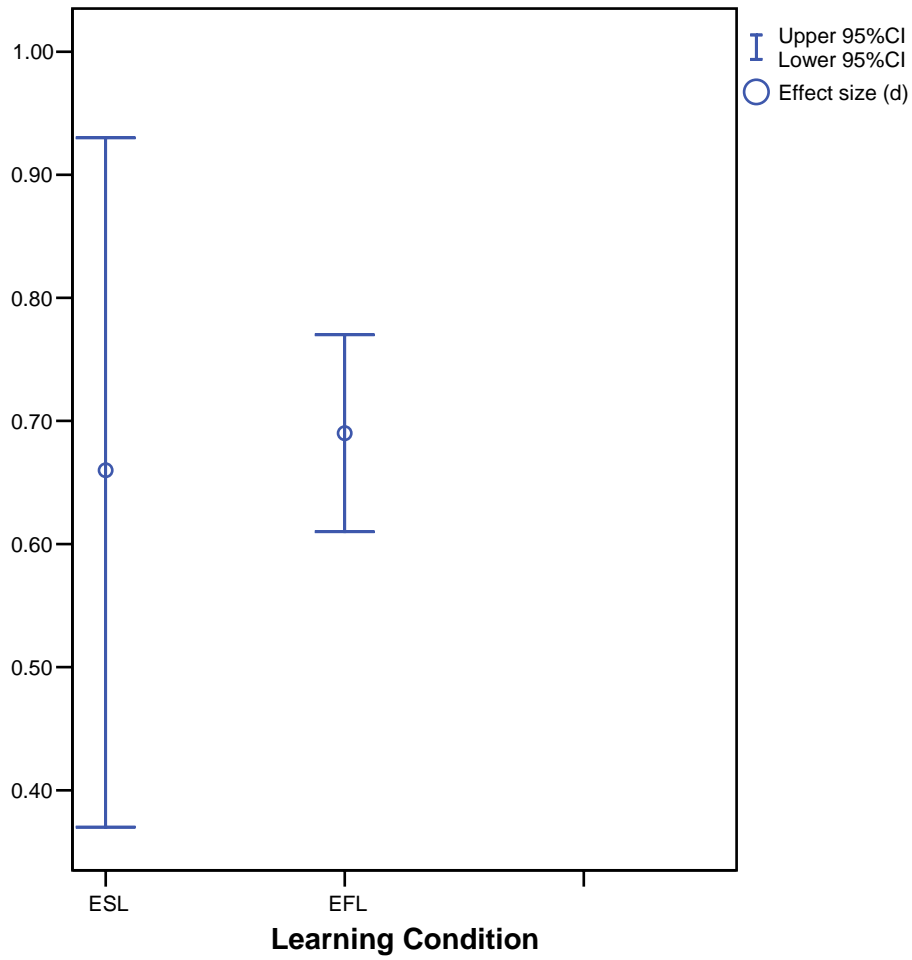


The wider confidence interval gives more confidence at the location of the population parameter. However, the shorter confidence intervals are more precise. Meta-analysis is weighted toward the more precise estimates. In Figure 4.2, decontextualizing instruction is both stronger and more precise. In other words, it is worth emphasizing that decontextualizing instruction yielded higher effects than the other two instructional approaches.

### ***3. Effects of Learning Conditions (ESL vs. EFL)***

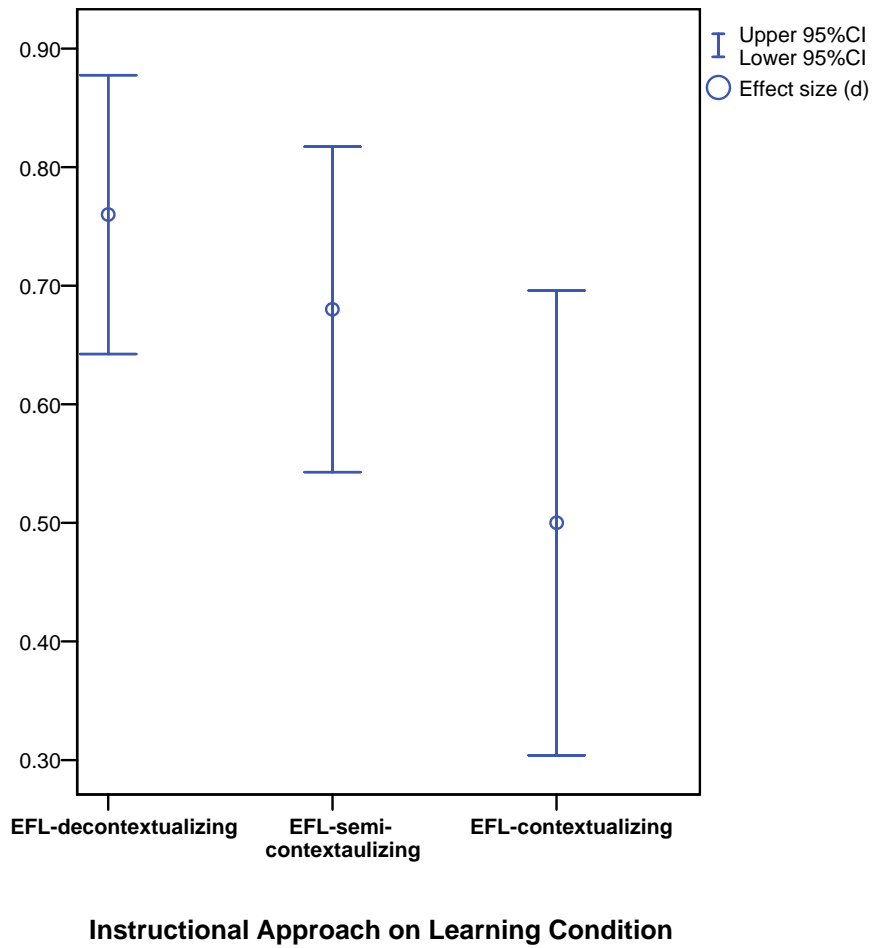
The effect size of the all studies conducted in the ESL conditions was  $d=.66$  (SE=.14) while the effect size of EFL studies was  $d=.69$  (SE= .04). Due to the small number of studies included in ESL (n=6) compared to EFL (n=35), this result of comparison should be interpreted with caution if there is a difference. Figure 4.3 shows that there is no significant difference at all between the two groups. It can be concluded that learning condition does not yield any different effect on vocabulary instruction.

To find the factors related to the effectiveness of instruction, further analysis was conducted adding the effects of three instructional approaches on EFL conditions. The contextualizing group in EFL has a little smaller effect size than the overall effect size of contextualizing instruction. Figure 4.4 displays the comparison of the effects of three types of vocabulary instruction on EFL conditions. However, their effect ranges still overlapped; it can be concluded that there is no statistically significant difference even though their mean effect sizes are quite different. The mean effect of decontextualizing instruction was  $d=.76$ , which is close to a large effect based on Cohen's (1977) guideline while the mean effect size of  $d=.50$  on the EFL-contextualizing group was considered as a medium effect.



	<i>d</i>	<i>SE</i>	<i>k</i>	<i>n</i>	<i>95% lower</i>	<i>95% upper</i>
ESL	.66	.14	6	234	.37	.93
EFL	.69	.04	35	2668	.61	.77

**Figure 4.3. Effects of learning conditions**



	<i>d</i>	<i>SE</i>	<i>k</i>	<i>n</i>	<i>95% lower</i>	<i>95% upper</i>
EFL-decontextualizing	.76	.06	12	1303	.64	.88
EFL-semi-contextualizing	.68	.07	15	943	.54	.82
EFL-contextualizing	.50	.10	8	422	.30	.70

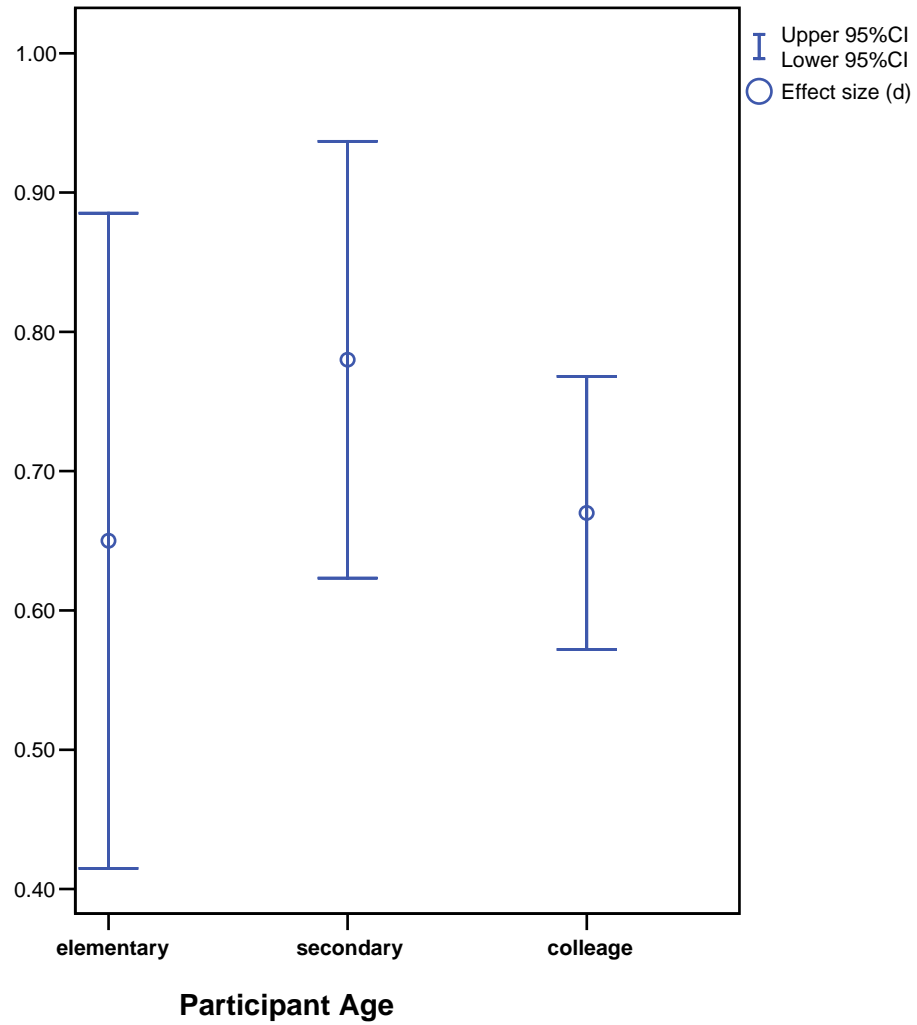
**Figure 4.4. Effects of instructional approaches on leaning condition**

It is also important to point out that the lower confidence interval  $d=.64$  on EFL-decontextualizing instruction and  $d=.54$  on EFL-semi-contextualizing instruction were both higher than the mean effect size of contextualizing instruction for EFL. It might be interpreted that the worst case of decontextualizing or semi-contextualizing instruction still yields a better effect on the average effect of contextualizing instruction especially in EFL conditions.

#### ***4. Effects of Participant Age***

Participant ages were other aspects that were investigated in this meta-analysis. From the between-group comparisons, it was found that 23 studies were conducted using university students or adult learners of English, 10 provided secondary student results, 5 provided elementary student results, and only 3 provided early childhood results. It is important to remember that from these descriptive statistics, it is clear that there were many more studies that reported on university students than the number of studies reporting on secondary or elementary combined.

Since there were only three observations available for the children as English learner category, it was impossible to conduct a meta-analysis for this group. Therefore, three group comparisons for elementary, secondary, and university students were made to determine the effectiveness of age variables. Studies with college student participants, i.e., the highest number of studies, showed the effect size  $d=.67$  ( $SE=.05$ ). The studies



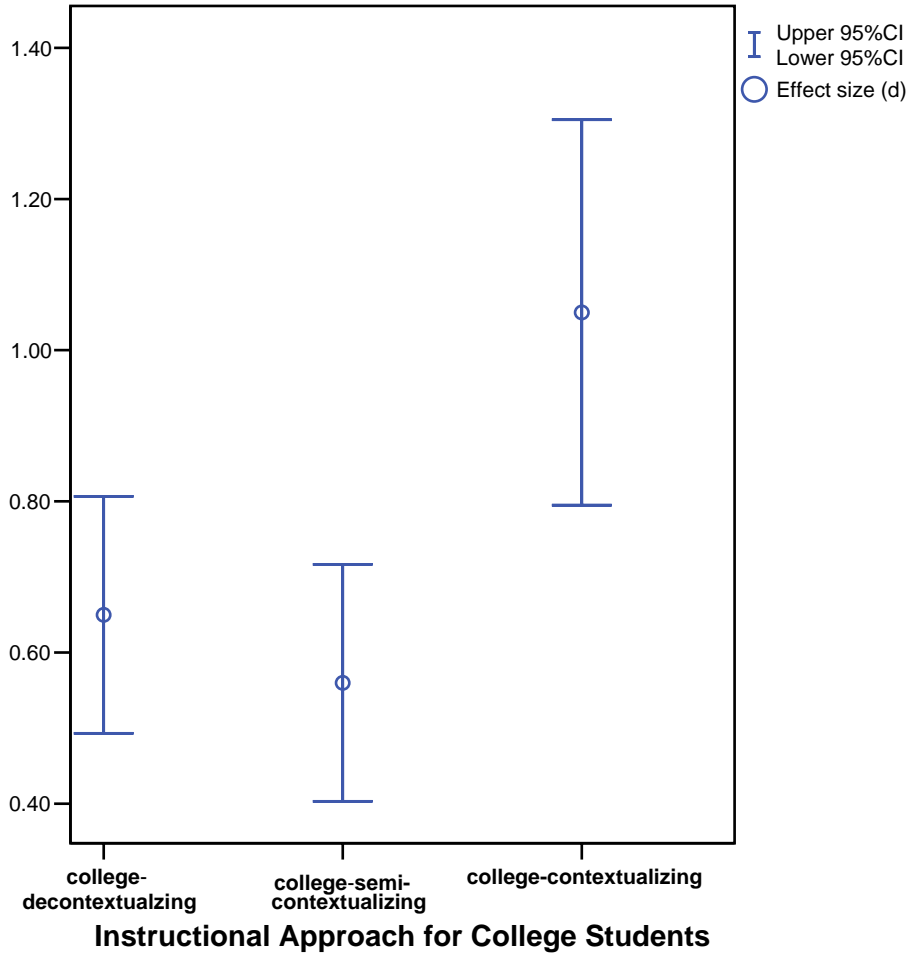
	<i>d</i>	<i>SE</i>	<i>k</i>	<i>n</i>	<i>95 % lower</i>	<i>95 % upper</i>
Elementary	.65	.12	5	310	.41	.89
Secondary	.78	.08	10	758	.62	.94
College	.67	.05	23	1738	.57	.77

\* Note: Children not included because of small number of studies

**Figure 4.5. Effects of participant age**

conducted for secondary students yielded a larger effect size  $d=.78$  ( $SE=.08$ ) while elementary student participants had a smaller one  $d=.65$  ( $SE=.12$ ). This result suggested that there were no statistically significant differences on this category. This is also displayed clearly in Figure 4.5.

For the further study, three instructional approaches conducting for university participants were meta-analyzed to determine the effects of each instruction on college students. The effect size of the contextual approach for college students yielded the very large  $d=1.05$  ( $SE=.13$ ) as opposed to the overall effect size of contextualizing instruction  $d=.58$ . To know the effects of decontextualizing and semi-contextualizing instruction on college students as well, the effect sizes of those groups were compared to the overall effect size of two instructional approaches. The effect size of decontextualizing group of college students was  $d=.65$  ( $SE=.08$ ), which is a lower effect than the overall effect of decontextualizing  $d=.76$ . For the effect of the semi-contextualizing for college students was  $d=.56$  ( $SE=.08$ ), which is also lower than the overall effect of semi-contextualizing  $d=.65$ . These results were homogeneous especially for the decontextualizing condition with chi-square of 2.89 even though the number of studies in this category was only five, but five is the minimum number to measure a chi-square. From the between-group comparisons, this category led to significant results. Comparisons yielded a medium effect for semi-contextualizing and decontextualizing groups, and a larger effect size for the contextualizing groups.



	<i>d</i>	<i>SE</i>	<i>k</i>	<i>n</i>	<i>95% lower</i>	<i>95% upper</i>
College-decontext	.65	.08	5	687	.49	.81
College-semi-context	.56	.08	13	779	.40	.72
College-context	1.05	.13	5	272	.80	1.30

**Figure 4.6. Effects of instructional approaches for college students**

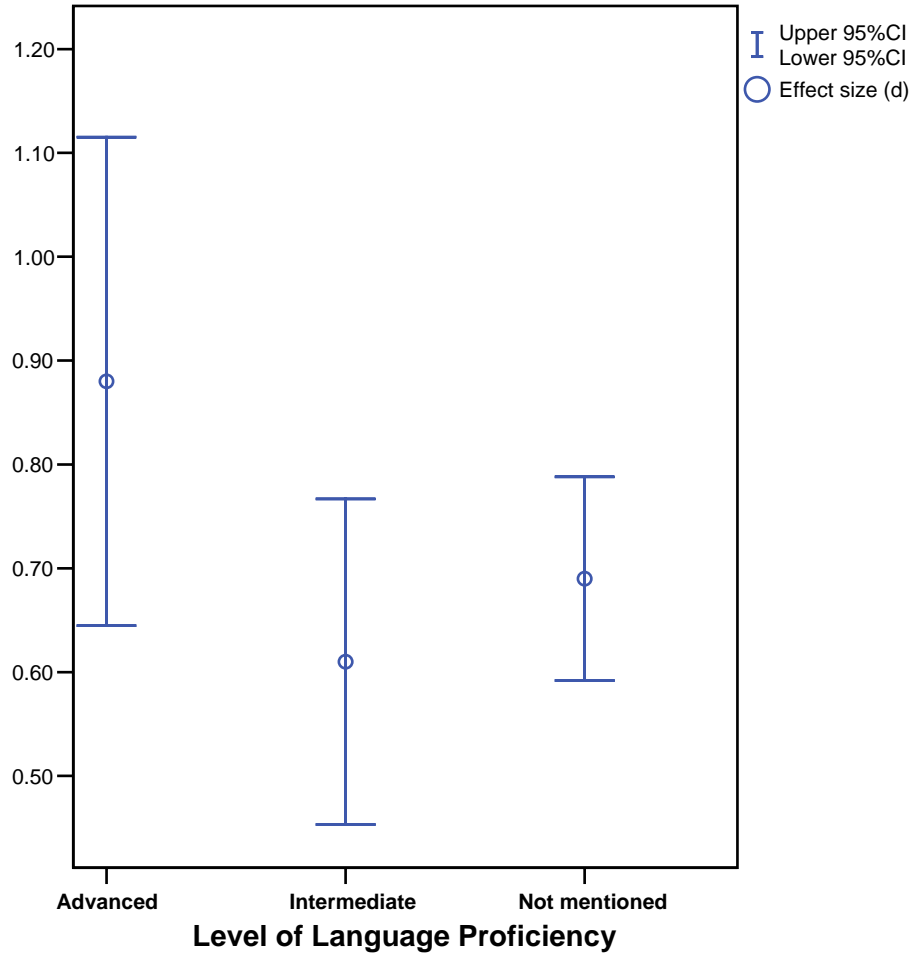
In other words, contextualizing vocabulary instruction for college students works better than the other two instructional approaches significantly. This is also shown clearly in Figure 4.6 with graphical presentation.

### ***5. Effects of Language Level***

Another investigation was made in order to explore the effects of the language level of English learners. Language level was defined at three different levels: beginner, intermediate, and advanced. However, there were only three studies for beginners. Therefore the results of this group were not included in this comparison. Also many studies did not report the levels of proficiency, which were considered a group. Numerical results for this category are presented in Figure 4.7 with a graph. The group of studies for advanced level of English learner yielded a larger effect size  $d=.88$  but with large variability ( $SE=.12$ ) across the six studies aggregated for this variables. The large margin of error related to advanced learners goes from a lower CI  $d=.65$  to an upper CI  $d=1.12$ . In addition, due to the small number of studies included in this group ( $k=6$ ), the result of this effect on advanced learners should be interpreted cautiously.

With a relatively large number of studies ( $k=11$ ) included in the group of intermediate learners, this group yielded a medium effect ( $d=.61$ ,  $SE=.08$ ). While the amount of error in the observation between the group of advanced and the group of intermediate is small, it is important to point out that the mean effect size of advanced learners  $d=.88$  is a large effect on learning vocabulary compared to a medium effect





	<i>d</i>	<i>SE</i>	<i>k</i>	<i>n</i>	<i>95% lower</i>	<i>95% upper</i>
Advanced	.88	.12	6	884	.64	1.12
Intermediate	.61	.08	11	640	.45	.77
Not reported	.69	.05	21	1828	.59	.79

\* Note: Beginners not included because of small number of studies

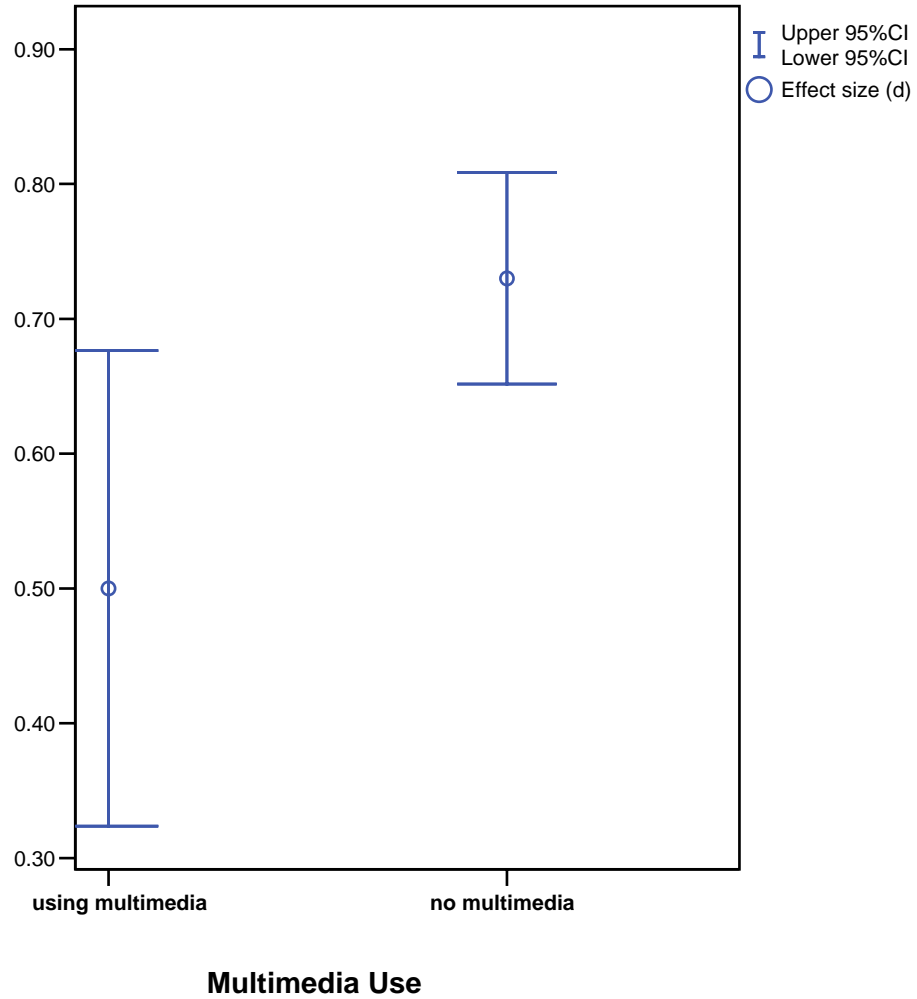
**Figure 4.7. Effects of the level of language proficiency**

$d=.61$  for intermediate learners. Therefore, it can be concluded that vocabulary instruction works better with advanced learners than with intermediate learners.

### ***6. Effects of Multimedia Use***

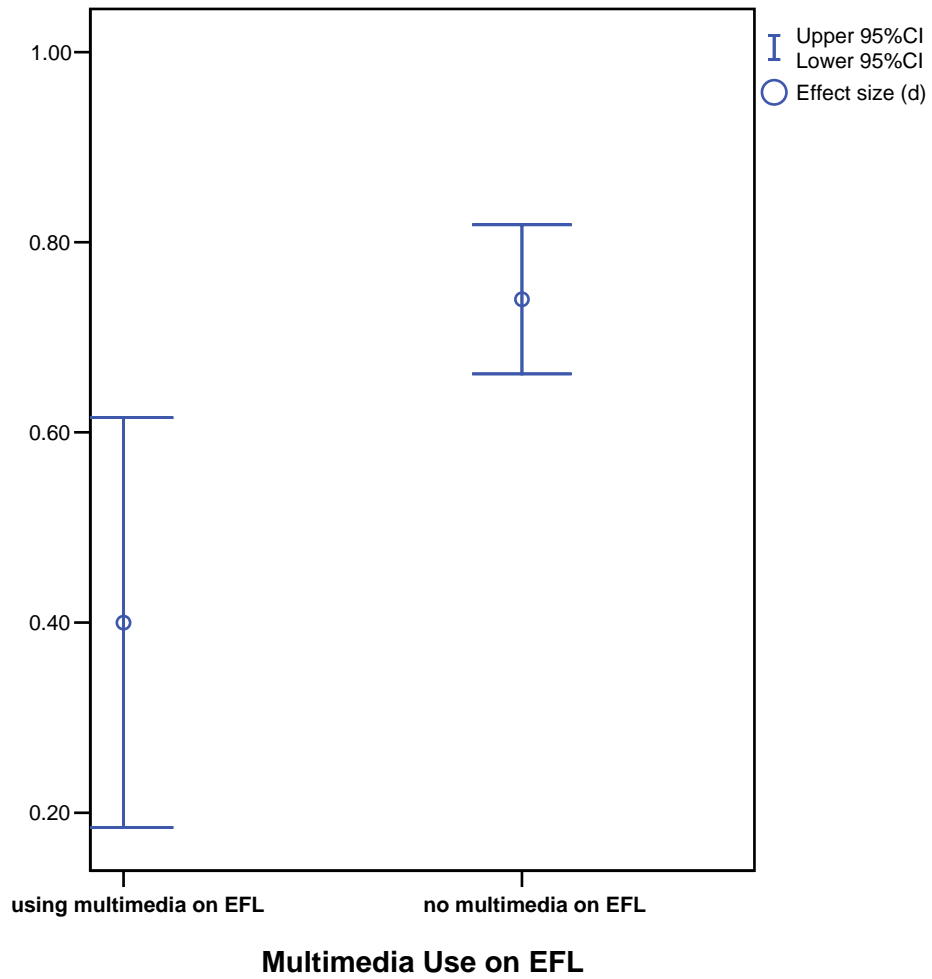
The effect of multimedia use was another variable to be investigated in this meta-analysis. The group of vocabulary instruction studies using computers as a primary teaching tool was compared with the group of studies conducted in classes without computer supports. Thirteen studies among 41 measured some kinds of instruction related to computer programs or using computers as an effective tool. Those are computer-assisted context embedded instruction (Kang, 1992), computer-supported learning of vocabulary items (Van Bussel, 1993), hypermedia-learning program (Al-Seghayer, 2001), computer-based electronic dictionary (Hill & Laufer, 2003, Laufer, 2005), vocabulary learning in multimedia context (Sun & Dong, 2004), computer-mediated negotiated interaction (Smith, 2004), and text, audio, and graphic aids in multimedia instruction (Kim, 2006).

Between-group comparison shows in Figure 4.8 that the group of studies without multimedia use yielded a larger effect size  $d=.73$  ( $SE=.04$ ) than the effect size for the multimedia using group ( $d=.50$ ,  $SE=.09$ ). Even though there is little overlap with their confidence interval, it could be concluded that the effect of multimedia use on vocabulary instruction did not work better than the instruction without multimedia.



	<i>d</i>	<i>SE</i>	<i>k</i>	<i>n</i>	<i>95% lower</i>	<i>95% upper</i>
Using multimedia	.50	.09	13	524	.32	.68
No multimedia	.73	.04	28	2378	.65	.81

**Figure 4.8. Effects of multimedia use**



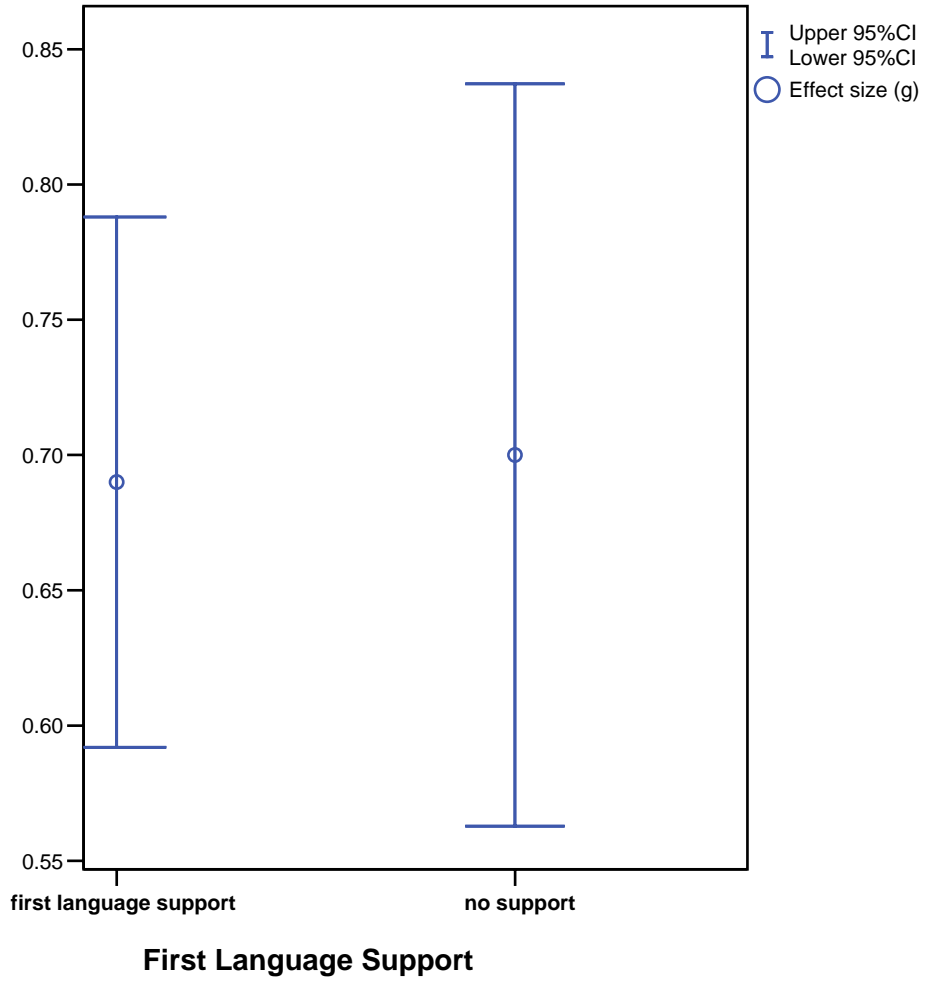
	<i>d</i>	<i>SE</i>	<i>k</i>	<i>n</i>	<i>95% lower</i>	<i>95% upper</i>
Using multimedia on EFL	.40	.11	7	398	.18	.62
No multimedia on EFL	.74	.04	28	2270	.66	.82

**Figure 4. 9. Effects of multimedia use on EFL condition**

For the further study, another computation was done only for the only EFL learners in order to determine the effectiveness of multimedia. The effect size of multimedia group on EFL conditions was  $d=.40$  ( $SE=.11$ ), which is smaller than the effect size of overall group. The effect size of no multimedia group was  $d=.74$  ( $SE=.04$ ) which is almost the same as the overall group. The margin of error ranges from the lower CI  $d=.18$  to an upper CI  $d=.62$  for the multimedia group. The low number of studies included in this group ( $k=7$ ) is likely responsible for this confidence interval and might not allow a trustworthy result compared to the result of no multimedia group. However, the worst case scenario of no multimedia group effect size  $d=.66$  is better than the upper  $d=.62$  of multimedia use group in the EFL condition. This is shown clearly with the graphical display in Figure 4.9. These findings support that vocabulary instruction without multimedia use are likely to results in better effects on learning vocabulary, especially for the EFL condition.

### ***7. Effects of First Language Supports***

Twenty four studies conducted under EFL conditions allowed for use of first language translations or bilingual dictionary look-up during experiments. The 17 studies including six studies conducted in ESL did not use first language support. From the between group comparison, it was found that there were no differences statistically. The effect size of first language support group was  $d=.69$  ( $SE=.05$ ), and the effect size of the no support group was  $d=.70$  ( $SE=.07$ ). Figure 4.10 displays the result of this category.



	<i>d</i>	<i>SE</i>	<i>k</i>	<i>n</i>	<i>95% lower</i>	<i>95% upper</i>
First language support	.69	.05	24	1925	.59	.79
No support	.70	.07	17	977	.56	.84

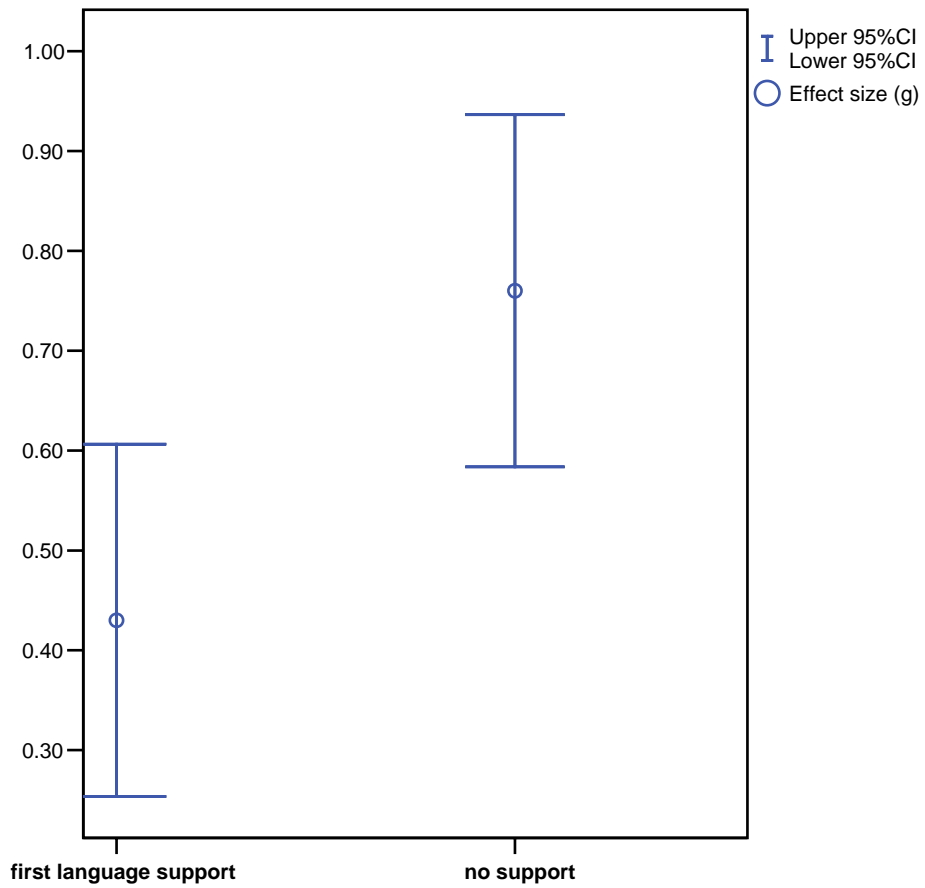
**Figure 4.10. Effects of first language supports**

For the further study, three instructional approaches were investigated with or without first language support in order to determine the effectiveness of first language supports.

Among the 12 studies of the decontextualizing group, nine studies allowed using translations or a bilingual dictionary; three studies did not have any of these supports. The overall effect size of decontextualized instruction was  $d=.76$  ( $SD=.06$ ) as mentioned above. The sub-group of decontextualizing with first language support yielded a little smaller effect size  $d=.73$  ( $SE=.06$ ) while the sub-group of without support did a larger  $d=.88$  ( $SE=.14$ ). Because there were not enough studies for no support group to conduct a meta-analysis, this effect may not be trustworthy.

Another computation was made for the semi-contextualizing and contextualizing categories as well. From the between group comparison results, it was found that the effect of first language support was significantly different between groups. The effect size of semi-contextualizing group with first language support was  $d=.43$  ( $SE=.09$ ) while the group without support was  $d=.76$  ( $SE=.09$ ). Since their confidence intervals overlap little between two groups in Figure 4.11, it could be concluded that the semi-contextualizing approach works better without first language support for learning new vocabulary.

However, the effect of first language support was the opposite on contextualizing instruction. The effect size of contextualizing group with first language support was  $d=.80$  ( $SE=.13$ ) while the effect size of the group without first language support was  $d=.25$  ( $SE=.16$ ). Findings show that the group of contextualizing instruction with first language support yielded a large effect with less variability compared to the other group without first language support, which had a small effect size with larger variability.

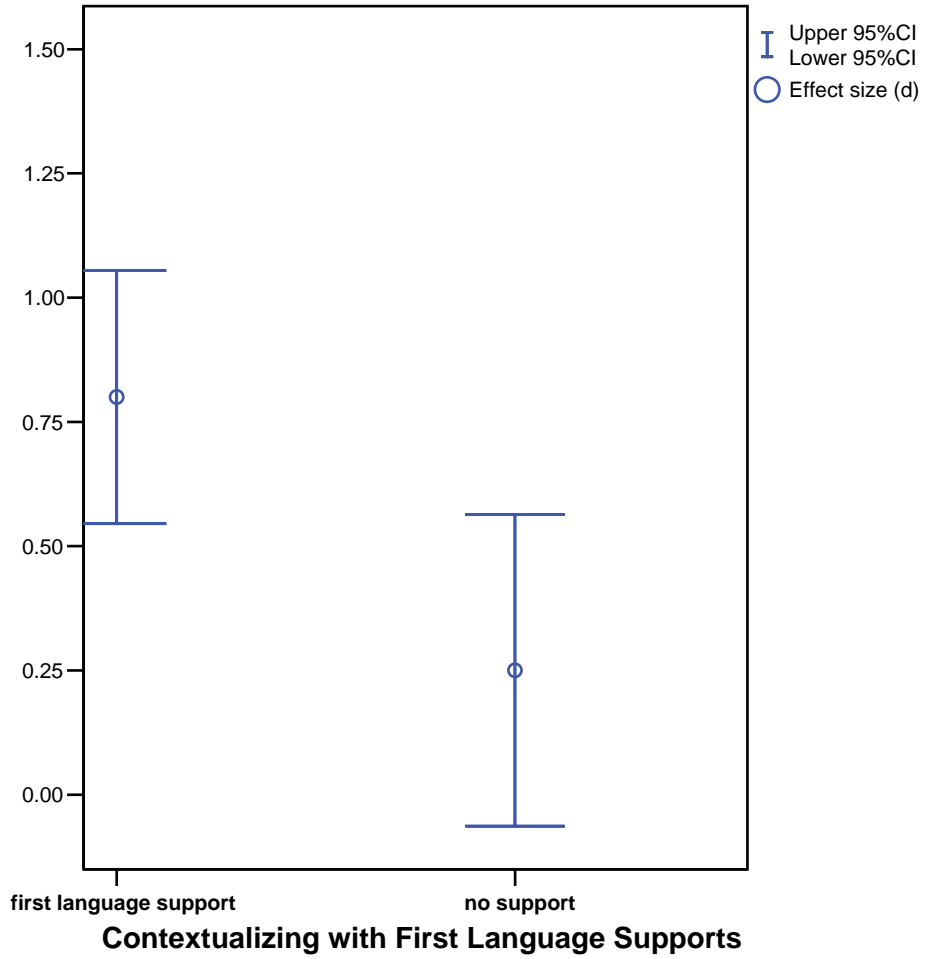


**Semi-contextualizing with First Language Support**

	<i>d</i>	<i>SE</i>	<i>k</i>	<i>n</i>	<i>95% lower</i>	<i>95% upper</i>
First language support.	.43	.09	9	534	.25	.61
No support	.76	.09	10	551	.58	.94

**Figure 4.11 Effects of first language supports on semi-contextualized instruction**





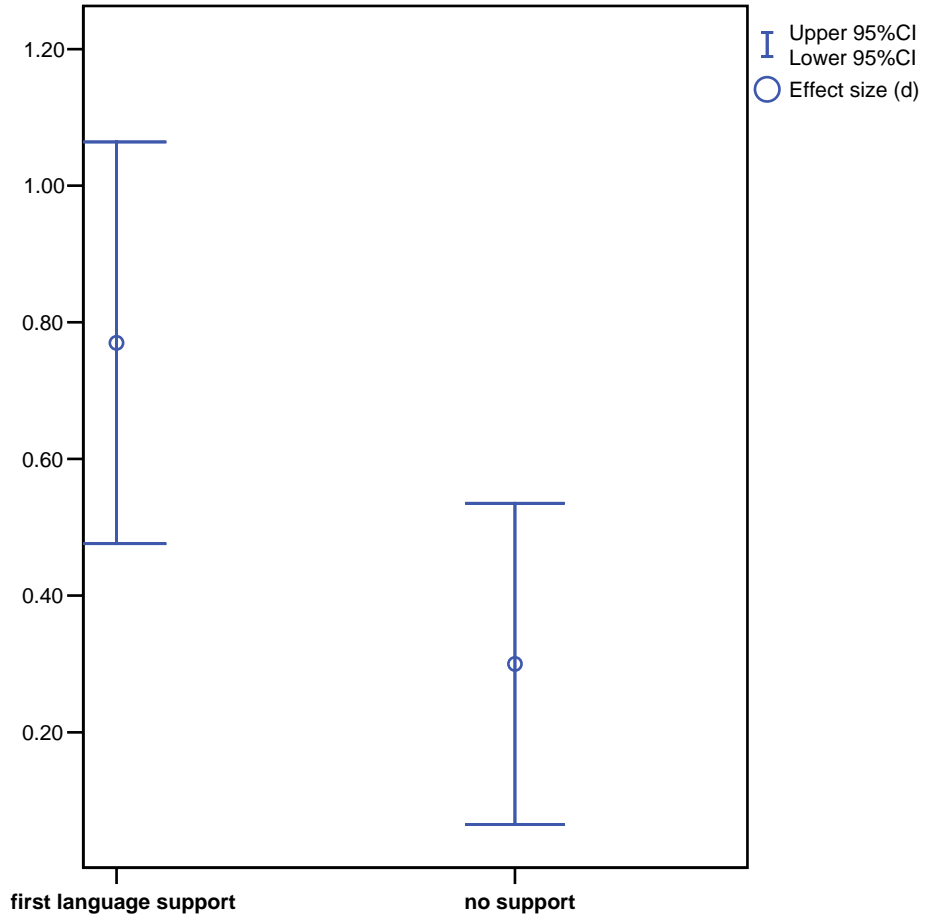
	<i>d</i>	<i>SE</i>	<i>k</i>	<i>n</i>	<i>95% lower</i>	<i>95% upper</i>
First language support	.80	.13	5	271	.55	1.05
No support	.25	.16	4	183	-.06	.56

**Figure 4. 12. Effects of first language supports on contextualized instruction**

Figure 4.12 presents the results from between group comparisons of first language support on contextualizing instruction.

The small number of studies ( $k=5$  and  $k=4$ ) might cause larger confidence interval. However, the effect of the worst case with first language support works better than the best case of the group without first language support on contextualizing instruction. It could be concluded that contextualized instruction is more effective with first language support.

The effect of first language support was investigated in the relationship of multimedia use as well. The overall effect size of multimedia use was  $d=.50$  ( $SE=.09$ ), which means a medium effect according to Cohen's (1977) guideline. Among 13 studies of using multimedia, four studies were conducted with first language supports while six studies using multimedia were conducted without first language supports. The effect size of first group was  $d=.77$  ( $SE=.15$ ), which means almost a large effect on vocabulary learning. For the second group without first language supports, the effect size was  $d=.30$  ( $SE=.12$ ), which is close to a small effect. Even though the number of studies is small, Figure 4.13 displays that there is a difference between the two groups. In other words, it can be concluded when vocabulary instruction use a multimedia as a tool, first language supports help learning new words better.



**First Language Supports on Multimedia Use**

	<i>d</i>	<i>SE</i>	<i>k</i>	<i>n</i>	<i>95 % lower</i>	<i>95 % upper</i>
First language support	.77	.15	4	222	.48	1.06
No support	.30	.12	6	302	.06	.54

**Figure 4. 13. Effects of first language supports on multimedia use**

## **CHAPTER V**

### **DISCUSSION, CONCLUSIONS AND RECOMMENDATION**

This chapter discusses and summarizes the main findings from the present meta-analysis, and provides recommendations for future research. The first section summarized the finding and draws conclusions directly connected to the research questions. General conclusions are presented in the second section. The last section discusses some recommendation for future research.

#### **Summary of Main Findings**

This meta-analysis used the summary results of 43 studies in order to determine the general state of vocabulary instruction studies. The main goals of this study included the categorization of the instructional methods, an examination of the relative effectiveness of these methods, and an examination of factors that may have affected the outcomes of studies.

*Question 1. How can vocabulary instruction used in quasi-experimental or experimental studies be categorized?*

The categorization of vocabulary instructional studies indicated the variety of instructional approaches studied in the past three decades. As indicated in Chapter IV and tables in Appendix B, studies were categorized as contextualized instruction, semi-contextualized instruction, or decontextualized instruction. The contextualizing category

included the studies focusing on reading and listening or speaking and writing practice as the main instruction. As opposed to contextualizing instruction, the decontextualizing category focused on word meaning separately. Flashcards and word lists, form-focused tasks, and using the dictionary were used in this approach. Other studies identified as semi-contextualizing category used the keyword method, semantic mapping, word or concept association, visual imagery, and aural imagery as the main instruction.

The three categories of studies were also divided by multimedia use and non-multimedia. The category of studies using multimedia as a teaching tool measured computer-related learning word programs. These computer programs are subcategorized as to whether they used first language supports in the programs.

*2. To what degree are these methods of instructions effective? What are the effects of such methods on learning vocabulary for English language learners' achievement?*

The overall effect size was  $d=.69$  for this meta-analysis. Since the effect size is the same as a  $z$  score on a normal curve, an effect size  $d=.7$  is the same as .2580 (about 26%). This means that if the control groups were to receive the treatment, their scores should improve about 26% on average. In other words, with practical meaning among about 1450 students (the number of students in control groups) who received the treatment, 375 students (multiply by .2580) would have been affected positively.

The highest effect size of the instructional categories was decontextualized instruction, the effect size  $d=.76$  which is considered close to a large effect by Cohen's (1977) guideline. The lowest instructional category was contextualized instruction, the

effect size  $d=.57$  which is lower than the overall effect size, but still considered a medium effect. The effect size of semi-contextualized instruction was  $d=.65$  which is close to the overall effect size of this study.

The effect size of using multimedia category was  $d= .50$ , which is a medium effect, but a lower effect than the effect size of no multimedia use. The effect size of studies without multimedia use was  $d=.73$ , which is a little higher than the overall effect.

However, the instructional approach with or without the first language supports does not appear to be different unless they are associated with different factors of vocabulary instruction. The effect sizes of both groups are the same as the overall effect size. For example, the effect size of first language supports was  $d=.69$ , but if associated with contextualizing instruction, the effect size went up to  $d=.80$ . It dropped down to  $d=.25$  without first language supports on contextualizing instruction.

*3. Under what conditions are these kinds of instruction effective? That is, what are the mediating effects of substantive variables affecting vocabulary instruction, such as the age of samples, their levels of English as a second language, geographical location, and the publication source of study?*

These three types of vocabulary instruction, contextualizing, semi-contextualizing and decontextualizing categories, are effective statistically whether the studies were conducted in ESL or EFL. Also these instructional approaches are effective regardless of whether participants are in elementary, secondary, and university. However, contextualized instruction for college students yielded larger effect size ( $d=1.05$ ) while

the effect size of semi-contextualized ( $d=.56$ ) and decontextualized instruction ( $d=.65$ ) for college students yielded medium effects on vocabulary learning.

The level of language proficiency was a factor to mediate the effect size. The effect size of advanced learners of English was  $d=.88$  which means above a large effect compared to the effect size of  $d=.61$  for intermediate learners. However, it is impossible to explore the relationship between the level of proficiency and the three types of instruction since there were not enough studies to conduct mediating effects in this case.

The instructional approach using multimedia yielded a smaller effect size ( $d=.50$ ) than the effect ( $d=.73$ ) of the studies without multimedia use. The effect size of studies with multimedia use dropped when the studies were conducted in EFL ( $d=.40$ ). However, when the studies with multimedia use was supported by first language translation or a electronic bilingual dictionary, the effect size went up ( $d=.77$ ), which is significantly different from the effect size of no first language supports on multimedia environments ( $d=.30$ ).

There were no different effects of using first language supports on vocabulary learning statistically in this study. However, the semi-contextualizing category worked more effectively without first language supports ( $d=.76$ ). As opposed to the above, some studies with first language supports were more effective on the contextualizing category of vocabulary instruction ( $d=.80$ ).

The publication sources of this study mediated the effect sizes as well. The mean effect size for twenty nine studies from journals was  $d=.80$  while it was  $d=.40$  in twelve studies from dissertations. The results found in journal articles are more positive than the results from dissertations. Because the study effects are somewhat higher in journal

articles than in dissertations, my study results might be concluded that these caused higher effect sizes than real.

## **Discussions and Conclusions**

The main finding of this study is that vocabulary instruction is effective overall for ESL and EFL learners although there are some instructional approaches which are not related to effective learning. It confirmed similar results of previous research of first language vocabulary learning, that any instructional approaches are better than nothing even though there is not a single best method of vocabulary instruction. (Petty, Herold, & Stol, 1968)

The NRP (2000) pointed out that conducting a formal meta-analysis could be difficult because there were a small number of research studies dealing with a relatively large number of variables. However, vocabulary research often appeared inconclusive, conflicting, or inapplicable. Conducting a meta-analysis with some selected variables was therefore helpful to compare the various instruction approaches.

### ***Effects of Instructional Approaches***

The current study suggests that vocabulary instruction for English language learners is effective. However, instruction must be appropriate to the learning conditions and learners' characteristics. Although most of the instruction is effective for English language learners, decontextualizing instruction is the most effective for them. Decontextualizing instruction is focusing on the word meaning separately from context



by using a flash card or word list, or any other activities that focus on forms or core meanings of words. It confirmed the results of many studies reviewed in Chapter II that favored to decontextualizing instruction, such as form-focused or word-focused instruction over contextualized instruction. For example, Ellis (2001) examined the role of form-focused instruction by reviewing 11 studies. He found that form-focused instruction can have a significant effect on the accuracy of use of grammatical structures and assist the acquisition of implicit knowledge.

Contextualizing instruction shows a lower effect size than the decontextualizing or semi-contextualizing instruction. That is even though contextualizing instruction is effective, learning is slow and not efficient for all of the cases, especially for ESL and EFL learners. Pickering (1982) found learning in context slightly more advantageous, but not enough to be considered as different achievements. Much of the research conducted in vocabulary instruction has failed to demonstrate any clear advantage of learning in contextualizing conditions (Nation, 1982). For instance, Coady (1993) concluded after exploring the basic argument for a mixed approach to vocabulary acquisition in ESL that the basic or core vocabulary should be taught, but the less frequent vocabulary would be learned better via context. However, even in that case, some techniques played an important role to learning effectively. Carter and McCarthy (1988) concluded that a mixture of approaches should be adopted since there are advantages and disadvantages between context-based inferential strategies and some other explicit vocabulary learning approaches such as key-word techniques, or translation in pairs, or using a monolingual or bilingual dictionary. In other words, the result of the current study is parallel with previous research mentioned above.

However, contextual instruction works better for the college students, which means that more advanced learners are able to learn new words during subconscious activities or reading. Huckin and Coady (1999) stated some advantages of incidental vocabulary learning (contextualizing) over direct introduction (decontextualizing). They concluded that the incidental learning within context is still seen as an important part of vocabulary building, especially among advanced learners. In addition, contextualizing instruction can be better if the instruction used first language supports together. That is, if students learned new words during reading or listening, the use of bilingual dictionaries or translation with their first languages might not be a waste of time and effort.

### ***Effects of Learning Conditions***

Since this study indicated that there are no different effects of vocabulary instruction on learning conditions between ESL and EFL, it could be concluded that vocabulary instruction is effective regardless of population. That is, the methods used for one population (ESL) also work for the second population (EFL). However, the unbalanced small number of studies conducted in ESL might cause an unreliable result in this category, even though seven studies in this category were enough to conduct a meta-analysis. Therefore, some questions remain in this category.

### ***Effects of Multimedia Use***

The use of multimedia in vocabulary instruction was found to be effective, but the effect size of the multimedia group was lower than the effect size of the group without multimedia use. Therefore, it could not be said that the use of computers is better than

classroom teachers. Computer programs are supposed to be a learning aid to facilitate the learning condition, but this result also supports the common conclusion that learning occurs better during the interaction with teachers and other students. However, if the computer programs are used with first language supports, they work better. The result of this meta-analysis provided statistical evidence. The effect size increased ( $d=.77$ ) when the computer programs are supported by first language translation or electronic bilingual dictionaries while the effect size of no first language supports on multimedia environments decreased ( $d= .30$ ).

### ***Effects of First Language Supports***

The use of primary language supports is a strategy widely recommended for second language acquisition; however, the research opposing the use of native language strategies was prevalent as well. Despite debating on the use of the primary language, that use can come in many forms, from direct instruction in the native language to translation of worksheets in the real world. Fraser's study (1999) showed that consulting a dictionary to confirm inference is a valuable strategy for lexical acquisition. Kroll and Curley (1988) stated that vocabulary learning in the beginning stage uses translation exclusively compared to vocabulary learning at the advanced levels. Indeed, in Sautermeister's (1989) study, he reported that vocabulary learning behavior among English learners at colleges. English language learners were not satisfied until they had found a first language equivalent to assist their learning.

Although there are debates on the effectiveness of first language use for second language learning (Fraser, 1999; Kroll & Curley, 1988; Sautermeister, 1989), the result of

this study yielded no differences with or without first language supports, unless the instruction was associated with the other specific factors. As mentioned Chapter IV, if the first language supports are added on the instruction of multimedia use, the effects are larger than supports without multimedia use. In addition, this study found that the first language supports are needed on contextualizing instruction to provide a more effective instruction.

### ***Implications for Classroom Teachers***

This meta-analysis confirms that vocabulary instruction for English language learners has some similarities and dissimilarities to that of first language learners. To determine effective instruction for English language learners, the results of this study are reviewed based on the implications of vocabulary instruction offered by the NRP (2000). Followings are the implications of vocabulary instruction for first language learners.

1. Vocabulary should be taught both directly and indirectly.
2. Repetition and multiple exposures to vocabulary items are important.
3. Learning in rich contexts is valuable for vocabulary learning.
4. Vocabulary tasks should be restructured when necessary
5. Vocabulary learning should entail active engagement in learning tasks.
6. Computer technology can be used to help teach vocabulary.
7. Vocabulary can be acquired through incidental learning.
8. How vocabulary is assessed and evaluated can have differential effects on instruction.

9. Dependence on a single vocabulary instruction method will not result in optimal learning

The first implication by the NRP indicated that vocabulary should be taught both directly and indirectly to native English learners. While this might also be true for English language learners, the results of this study provides more support for the notion that direct instruction is more beneficial for second language vocabulary learning. Second, according to the NRP, learning in rich contexts is valuable for first language vocabulary learning. Findings in this support the NRP position but only for adult learners. For secondary and elementary students, this study found that decontextualizing instruction is more effective than learning in rich contexts for the English language learners. Third, the NRP suggests that computer technology can be used to help teach first language vocabulary. This study found that this is true if the conditions are met with certain levels of computer use such as adding first language supports or with contextualizing instruction, but teachers or classroom activities are more helpful than computers generally.

Fourth, another implication by the NRP was that vocabulary can be acquired through incidental learning. However, this study found that incidental learning, an example of contextualizing instruction, is not as effective for English language learners as semi-contextualizing or decontextualizing instruction. To determine more effective instruction for English language learners, a consideration of the characteristics of learners and environments as well as instructional approach should be made. From this meta-analysis, it was concluded that effective instruction for English language learners depends on how vocabulary instruction is used with appropriate conditions. ESL and EFL

classroom teachers need to consider the differences of first and second language vocabulary acquisition and students' characteristics as learners.

### ***Limitations of the Study***

It is worth noticing that these conclusions are based on only the reports collected for meta-analysis. Even though the conclusions were derived as a generalization from 43 studies collected and examined, this generalization of vocabulary instruction for English language learners does not begin to address this entire collection of vocabulary research. Because of the limitations of meta-analysis, all the qualitative studies and the quantitative studies that do not fit into this study's criteria were left out from this review. Certainly, it can be concluded that the effect size of each group indicates the effectiveness of that instruction. However, the possibility exists that these studies can not represent the entire population of each group. Another disappointment was the unequal number of studies distributed to each group. Unbalanced numbers of some groups might cause untrustworthy results, and the small number of studies in some groups prevented me from conducting a further analysis for mediating effects.

In summary, 43 studies produced large positive effects for vocabulary instruction. Even with the problems mentioned above, the procedures of meta-analysis were followed to accomplish the answers to my research questions and with credible research.

## Recommendations

This section proposes several recommendations for future research. First, a research synthesis of studies that use qualitative methods is needed in order to complete an explanation of the effect of vocabulary instruction because much more research were already conducted with qualitative frameworks on vocabulary instruction for English language learners. In addition, due to the methodology of meta-analysis, the current study could not synthesize some important results of quantitative studies. For example, if the study does not report means and standard deviations, or if the study measured N=words, not N= participants, these studies could not be included in this meta-analysis. Also, if the study design did not have comparison groups, it was removed from this meta-analysis because of the statistical limitation. Therefore, a qualitative review of research for those excluded from this analysis is needed.

Second, more quantitative studies conducted with ESL students are needed in order to understand better how ESL learners may differ from EFL learners in learning English vocabulary effectively. The search in this meta-analysis resulted in seven ESL studies compared to 36 EFL studies for this meta-analysis. Several ESL vocabulary studies were found but many were qualitative.

Third, more than half of the studies in this meta-analysis were conducted with adult learners of English. One of the reasons why many studies were conducted with college learners might be easy access to the subjects. To better understand the effects of level of proficiency and the effects of age on vocabulary acquisition, more studies are needed with elementary, secondary students, as well as primary children.

Fourth, in order to generalize the conclusions of the effect of vocabulary instruction for second language learners, a meta-analysis needs to be conducted that includes second/foreign languages other than English. Indeed, there were many studies about vocabulary learning in Spanish as a second/foreign language, French as a second/foreign language, German as a second/foreign language and so on. Therefore, including those studies as data for meta-analysis will widen the general conclusion of vocabulary learning in the area of second language research.

Last, a slightly different meta-analysis also on vocabulary is needed for the future. Since meta-analysis on vocabulary instruction is still rare, the research coded from a different perspective will provide additional information about effective vocabulary learning. To capture better the effect of instruction, further research using meta-analysis as a tool and with new categories is needed.



## References

- Alexander, P. A., Schallert, D. L., & Hare, V. C. (1991). Coming to terms: How researchers in learning and literacy talk about knowledge. *Review of Educational Research*, 61, 315-343.
- Al-Seghayer, K. (2001). The effect of multimedia annotation modes on L2 vocabulary acquisition: A comparative study. *Language, Learning & Technology*, 5(1), 202-229.
- Anderson, R. C. & Freebody, P. (1983). Reading comprehension and the assessment and acquisition of word knowledge. In B. Hutton (Ed.), *Advances in reading/language research: A research annual*. Greenwich, CT: JAI Press.
- Barcroft, J. (2004). Second language vocabulary acquisition: a lexical input processing approach. *Foreign Language Annals*. Vol.37.No.2. 200-208.
- Beck, I. L., Perfetti, C. A., & McKeown, M. G. (1982). The effects of long-term vocabulary instruction on lexical access and reading comprehension. *Journal of Educational Psychology*, 74, 506-521.
- Beck, I.L. & McKeown, M. G. (1991). Conditions of vocabulary acquisition. In R. Barr, M.L. Kamil, P.B. Mosenthal, and P.D. Pearson (Eds.) *Handbook of reading research* (Vol.II, pp.789-814). New York: Longman.
- Blachowicz, C. & Fisher, P (2002). Vocabulary Instruction. In M, Kamil, P, B, Mosenthal, P.D. Pearson, and R. Barr (Eds.) *Handbook of reading research volume III*. (pp 269-284) Mahon, NY: Lawrence Erlbaum Associates.
- Bot, K., Paribakht, T. S. & Wesche, M. B. (1999). Toward a lexical processing model for the study of second language vocabulary acquisition. *Studies in second language acquisition*, 19. 309-329.
- Brown, C., Sagers, S., & LaPorte, C. (1999). Incidental vocabulary acquisition from oral and written dialogue journals. *Studies in second language acquisition*, 21, 259-283.
- Carter, R. & McCarthy, M. (1988). *Vocabulary and language teaching*. London: London.
- Chall, J. (1987). Two vocabularies for reading: Recognition and meaning. In M. McKeown and M. Curtis (Eds), *The nature of vocabulary acquisition*. (pp.7-18). Hillsdale, NJ: Lawrence Erlbaum Associates.

- Coady, J. (1993). Research on ESL/EFL vocabulary acquisition: putting it in context. In T. Huckin, M. Haynes, and J. Coady (Eds.) *Second language learning and vocabulary learning*. Norwood, NJ : Ablex Publishing Corporation.
- Coady, J., Magoto, J., Hubbard, P., Graney, J., & Mokhtari, K. (1993). High frequency vocabulary and reading proficiency in ESL readers. In T. Huckin, M. Haynes, and J. Coady (Eds), *Second language reading and vocabulary learning* (pp217-228). Norwood, NJ: Ablex.
- Cobb, T. & Horst, M. (2001). Growing academic vocabulary with a collaborative on-line database. In B. Morrison, D. Gardner, K. Keobke, and M. Spratt (eds.) *ELT Perspectives on Information Technology & Multimedia: Selected Papers from the ITMELT 2001 Conference 1st & 2nd. English Language Centre, Hong Kong Polytechnic University*, pp. 189-226.
- Cohen, J. (1977). *Statistical power analysis for the behavioral sciences*. New York: Academic Press.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*(2<sup>nd</sup>). Hillsdale, NJ: Erlbaum.
- Cook, T. D. & Leviton, L.C. (1980). Reviewing the literature: a comparison of traditional methods with meta-analysis. *Journal of Pers*, 48, 449-472.
- Cooper, H. & Hedge, L.V. (Eds). (1994). *The handbook of research synthesis*. New York: Ressel Sage Foundation.
- Dale, E. (1965). Vocabulary measurement: Techniques and major findings. *Elementary English*, 42, 895-901.
- DeCarrico, J.S (2001). Vocabulary learning and Teaching. In M.C. Murcia (Eds.) *Teaching English as a second language*. Boston, MA : Heinle and Heinle
- Ellis, R. (2004). Does form-focused instruction affect the acquisition of implicit knowledge? : a review of the research. *Studies in second language acquisition*, 24, 223-236.
- Fan, M. Y.(2003). Frequency of use, perceived usefulness, and actual usefulness of second language vocabulary strategies: A study of Hong Kong learner. *The modern language journal*, 87, 222-241.
- Fitzgerald, J. (1995). English-as-a-second –language learner’s cognitive reading process: a review of research in the United States. *Review of educational research*, Vol.65, 145-190.

- Folse, K.S. (2004). *Vocabulary myths; applying second language research to classroom teaching*. Ann arbor, MI: the university of Michigan press.
- Fraser, C. (1999). Lexical processing strategy and vocabulary learning through reading. *Studies in Second language Acquisition*, 21, 225-241.
- Fukkink, R.G. & de Glopper, K. (1998). Effects of instruction deriving word meaning from context: a meta-analysis. *Review of Educational Research*, 68 (4), 450-69.
- Gaudio, V. (2003). *Improving Reading Skills in ESL Students through an Intensive Vocabulary Building Program*. Master of Arts Action Research Project, Saint Xavier University and SkyLight Professional Development Field-Based Master's Program(ERIC Document Reproduction Service No. ED 479899).
- Glass, G. V. (1976). Primary, secondary and meta-analysis of research. *Educational Researcher*, 5, 308
- Glass, G.V., McGaw, B., & Smith, M.L. (1981). *Meta-analysis in social research*. Beverly Hills, CA: Sage.
- Green, B. F., and Hall, J. A. (1984). Quantitative methods for literature reviews, *Annual Review of Psychology*, 35, 37-53
- Green, J. & Oxford, R. (1995). A closer look at learning strategies, L2 proficiency, and gender. *TESOL quarterly*, 29, 261-97.
- Groot, A.& Keijzer, R. (2000). What is hard to learn is easy to forget : The roles of word concreteness, cognate status, and word frequency in foreign-language vocabulary learning and forgetting. *Language Learning*, 50 (1), 1-56.
- Gu, Y. & Johnson, R.K. (1996). Vocabulary learning strategies and language learning outcomes. *Language Learning*, 49, 643-679.
- Hansen, L., Umeda, Y. & McKinney, M. (2002). Savings in the relearning of second language vocabulary: the effects of time and proficiency. *Language learning*, 52, 653-678.
- Harris, M. J. & Rosenthal, R. (1986). Mediating of interpersonal expectancy effect: 31 meta-analyses. *Psychological Bulletin*. 97, 363-386.
- Hedge, L.V. (1982). Fitting categorical models to effect sizes from a series of experiments. *Journal of Psychological Bulletin*, 92, 490-499.
- Hedge, L.V. & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando, FL: Academic Press.

- Hill, M. (2000). English vocabulary: tests and tasks. (ERIC Document Reproduction Service No. ED 462001)
- Hill, M. & Laufer, B. (2003). Type of task, time-on-task and electronic dictionaries in incidental vocabulary acquisition. *IRAL*, 41, 87-106.
- Hu, C. (2003). Phonological memory, phonological awareness, and foreign language word learning. *Language Learning*, 53(3), 429-462.
- Huckin, T. & Bloch, J. (1993). Strategies for inferring word-meanings in context: A cognitive model. In T. Huckin, M. Haynes, and J. Coady (Eds), *Second language and reading and vocabulary learning* (pp. 153-178). Norwood, NJ: Ablex.
- Huckin, T. & Coady, J (1999). Incidental vocabulary acquisition in a second language: A Review. *Studies in second language acquisition*, 21, 181-193.
- Hulstijn, J. H. (1992). Retention of inferred and given word meaning: Experiments in incidental vocabulary learning. In P.J.L. Arnaud & H. Bejoint (Eds.), *Vocabulary and applied linguistics* (pp.113-125). London: Macmillan.
- Hulstijn, J.H., Hollander, M., & Greidanus, T. (1996). Incidental vocabulary learning by advanced foreign language students: The influence of marginal glosses, dictionary use, and reoccurrence of unknown words. *Modern Language Journal*, 80(3), 327-339.
- Hulstijn, J. H. & Laufer, B (2001) Some Empirical Evidence for the involvement load hypothesis in Vocabulary acquisition, *Language Learning*, 51(3), 539-558.
- Hunt, M. (1997). *How science takes stock: The story of meta-analysis*. NY: Russell Sage Foundation.
- Hunter, J.E., Schmidt, F. L., & Jackson, G.B. (1982). *Meta-analysis: cumulating research findings across studies*. Newberry Park, CA: SAGE Publications.
- Hymes, D. (1972). On communication competence. In J.P. Pride and J. Holmes (Eds) *Sociolinguistics: Selected reading*. Harmondsworth, UK: Penguin.
- Irvin, J.L. (1990). Vocabulary knowledge: Guidelines for instruction. Washington, DC: National Education Association.
- Jiang, N. (2004). Semantic transfer and its implications for vocabulary teaching in a second language. *The Modern Language Journal*, 88 (3), 416-432.

- Kojic-Sabo, I. & Lightbrown, P. (1999). Students' approaches to vocabulary learning and their relationship to success. *The Modern Language Journal*, 83, 176-192.
- Krashen, S. (1989). We acquire vocabulary and spelling by reading: Additional evidence for the input hypothesis. *The Modern Language Journal*, Vol. 73, 440-464.
- Kulik, J. A. & Kulik, C. L. (1989). Meta-analysis in education. *International Journal of Educational Research*, 13, 221-340.
- Laufer, B (2003). Vocabulary acquisition in a second language: Do learners really acquire most vocabulary by reading some empirical evidence. *The Canadian Modern Language Review*. 59 (4), 567-587.
- Laufer, B. (2005). Focus on form in second language vocabulary learning. *EUROSLA Yearbook*, 5, 223-250.
- Laufer, B & Goldstein, Z. (2004) Testing Vocabulary knowledge: Size, Strength, and Computer Adaptiveness. *Language Learning* 54(3), 399-436.
- Laufer, B & Hill, M. (2000). What lexical information do L2 learners select in a CALL dictionary and how does it affect word retention? *Language learning & technology*, 3(2), 58-76.
- Lazaraton, A. (2004). Gesture and speech in the vocabulary explanations of one ESL teacher: A microanalytic inquiry. *Language learning*, 54(1), 79-117.
- Lessard-Clouston, M. (1996). ESL Vocabulary Learning in a TOEFL Preparation Class: A Case Study. *Canadian Modern Language Review*, 53 (1), 97-119.
- Lipsey, M. & Wilson, B (2001). *Practical meta-analysis*. Sage Publication, Thousand Oaks, CA.
- Mezynski, K. (1983). Issues concerning the acquisition of knowledge: effects of vocabulary training on reading comprehension. *Review of Educational Research*, Vol.53, 253-279.
- Nagy, W. E. & Scott, J. A. (2000). Vocabulary process. In M, Kamil, P, B, Mosenthal, P.D. Pearson, and R. Barr (Eds.) *Handbook of reading research volume III*. (pp 269-284) Mahon, NY: Lawrence Erlbaum Associates.
- Nagy, W. E. & Scott, J. A. (2006). The state of vocabulary research in the mid-1980s. In K. A. Dougherty Stahl and M. C. McKenna (Eds.) *Seeking understanding in how*

*to teach reading: Selected works by Steven Stahl and a memorial Festschrift.*  
New York: Guilford Publishing.

- Nation, I.S.P. (1990). *Teaching and learning vocabulary.* Boston: Heinleand Heinle.
- National Reading Panel. (2000). *Report of the teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction.* ( NIH Pub. No. 00-4754). Washington, DC: U.S. Government Printing Office.
- Oxford, R. & Crookall, D. (1990). Vocabulary learning: A critical analysis of techniques. *TESL Canada Journal*, 7 P9-30.
- Paribakht, T. & Wesche, M. (1999). Reading and “incidental” L2 vocabulary acquisition:an introspective study of lexical inferencing. *Studies in second language acquisition*, 21, 195-224.
- Paribakht, T. & Wesche, M. (1993). The relationship between reading comprehension and second language development in a comprehension-based ESL program. *TESL Canada journal*, 11(1), 9-29.
- Petty, W., Herold, C., & Stoll, E. (1968). *The state of the knowledge about the teaching of vocabulary.* Cooperative Research Project No. 3128. Champlain, IL: National Council of teachers of English.
- Prince, P. (1996). Second language vocabulary learning: the role of context versus translations as a function of proficiency. *The Modern Language Journal*, 80, 478-493.
- Qian, D.D. (1996). ESL Vocabulary Acquisition: Contextualization and decontextualization. *Canadian Modern Language Review*, 53, 120-142.
- Quin, D. D. (2002). Investigation the relationship between vocabulary knowledge and academic reading performance: An assessment perspective. *Language learning*, 52(3), 513-536.
- Raptis, H. (1997). Is second Language reading vocabulary best learned by reading? *Canadian Modern Language Review Vol. 53*, 566-580.
- Rodriguez, M. & Sadoski, M. (2000). Effects of rote, context, keyword, and context/keyword methods on retention of vocabulary in EFL classroom. *Language Learning*, 50, 385-412.
- Rosental, R. (1980). *Quantitative assessment of research domains: New directions for methodology of social and behavioral science.* 5. San Francisco: Jossey-Bass.

- Rosental, R. (1984). *Meta-analytic procedures for social research*. Beverly Hills, CA : Sage.
- Rossi, P. & Wright, S. (1977). Evaluation research: An Assessment of theory, practice, and politics. *Evaluation Quarterly*, 1, 5-52.
- Rudner, L., Glass, G. V., Evartt, D. L., & Emery, P.J. (2000). *A user's guide to the meta-analysis of research studies*. ERIC Clearinghouse on Assessment and Evaluation, University of Maryland (2000).
- Sanaoui, R. (1995). Adult learners' approaches to learning vocabulary in second languages. *The Modern Language Journal*. 79, 15-28.
- Schmitt, N. (1998). Tracking the incremental acquisition of second language vocabulary: a longitudinal study. *Language Learning*, 48(2), 281-317.
- Schunk, H. (1999). The Effect of Singing Paired with Signing on Receptive Vocabulary Skills of Elementary ESL Students. *The Journal of Music Therapy*, 36 (2), 110-124.
- Smith, B. (2005). The relationship between negotiated interaction, learner uptake, and lexical acquisition in task-based computer-mediated communication. *TESOL quarterly*, 39 (1), 33-57.
- Smith, B. (2004). Computer-mediated negotiated interaction and lexical acquisition. *Studies in second language acquisition*, 26, 365-398.
- Stahl, S. A. (1983). Differential knowledge and reading comprehension. *Journal of Reading behavior*, 15, 33-50.
- Stahl, S. A. & Fairbanks, M. M. (1986). The effects of vocabulary instruction: a model-based meta-analysis. *Review of Educational Research*, Vol. 56, 72-110.
- Sternberg, R. (1987). Most vocabulary is learned from context. In McKeown and M. Curtis (Eds). *The nature of vocabulary acquisition* (pp.89-106). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Swanborn, M.S. & de Glopper, K. (1999). Incidental word learning while reading: a meta-analysis. *Review of Educational Research*, 69(3), 265-85.
- Treville, M. (1996). Lexical learning and reading in L2 at the beginner level: The advantage of cognates. *Canadian Modern Language Review*, 53, 173-190.

- Walberg, H. J. & Haertel, E. H. (1980). Research integration: the state of the art. *Special issue of Evaluation in Education: An International Review Service*, 4, 1-135.
- Wa-Mbaleka, S. (2006). A meta-analysis investigating the effects of reading on second language vocabulary learning (Doctoral dissertation, Northern Arizona University, 2003). *Dissertation Abstracts International*.
- Wesche, M. B. & Paribakht, T. S. (2000). Reading-based exercises in second language vocabulary learning: an introspective study. *The Modern Language Journal*, 84, 196-213.
- Wilson, D.B. (1995). *The role of method in treatment effect estimates: Evidence from psychological, behavioral, and educational treatment intervention meta analyses*. Doctoral dissertation, Claremont Graduate School, Claremont, CA.
- Wode, H. (1999). Incidental vocabulary acquisition in the foreign language classroom. *Studies in Second language Acquisition*, 21, 243-258.
- Wolf, F. (1986). *Meta-Analysis: Quantitative methods for research synthesis*. Sage Publication; Newbery park, CA.
- Wolter, B. (2001). Comparing the L1 and L2 mental lexicon: a depth of individual word knowledge model. *Studies in second language acquisition*, 23, 41-69.
- Zahar, R., Cobb, T., & Spada, N. (2001). Acquiring vocabulary through reading: Effect of frequency and contextual richness. *The Canadian Modern Language Review*, 57 (4), 541-572.
- Zareva, A., Schwanenflugel, P., & Nikolova, Y. (2005). Relationship between lexical competence and language proficiency. *SSLA*, 27, 567



## APPENDIX A

### Coding Sheet

1. studid: study ID number
2. ES: effect size
3. ReNo: record number
4. OE: observed effect
5. AU-author
6. Pubdt: publication date
7. Pubtype: publication type
  1. journal article
  2. dissertation
8. Firstlanguage
8. rg-region where research took place
  - 1=the United State/
  - 2=Canada
  - 3=Europe
  - 4=Asia
  - 5=other
9. cd: learning conditions
  1. ESL
  2. EFL
10. parage :participant age
  - 1=children
  - 2=elementary students
  - 3=secondary students
  - 4=college students
11. langlevel: participant language level
  1. beginners
  2. intermediates
  3. advanced learners
  9. not mentioned
12. IA :instructioanl approach
  - 1= contextualizing :incidental acquisition from reading (Reading and Listening Practice, Speaking and Writing Practice )
  - 2= Semi contextualizing( semantic elaboration, Word Groupings, Word or Concept Association, Visual Imagery, Aural Imagery, Keyword, Physical Response , Physical Sensation, Semantic Mapping )
  - 3= decontextualizing ( focus on form, word list, Flashcards, Dictionary Use,)
13. using multimedia ( mediating computer, audio, video)
  1. Yes
  2. No
14. using first language support( translation & dictionary)

1. Yes
2. No

15. Vote:vote count

- 1=significant neg.
- 2=non-significant neg
- 3=non-significant pos
- 4=significant pos

16. n1:sample size of experimental group
17. n2:sample size of control group
18. mean1: mean score of experimental group
19. mean2: mean score of control group
20. sd1: standard deviation of experimental group
21. sd2 : standard deviation of control group
22. ASINS : assignment of instruction
23. test

## APPENDIX B

```
SORT CASES BY  
  studid (A) .
```

```
compute trial=studid.  
execute.
```

```
compute poolvar=((n1-1)*(sd1**2)+(n2-1)*(sd2**2))/(n1+n2-2) .
```

```
compute s = sqrt(poolvar).
```

```
compute g = (mean1 - mean2) / s.  
execute.
```

```
*CREATES ONE MEAN(ES) PER STUDY.  
AGGREGATE  
  /OUTFILE='C:\aggrmijin.sav'  
  /BREAK=Studid  
  /g = MEAN(g)  
  /N_BREAK=N.
```

```
*MERGE FILES--REMEMBER TO SORT FIRST BY STUDY.  
MATCH FILES /FILE=*  
  /FILE='C:\aggrmijin.sav'  
  /RENAME g=g_final  
  /BY Studid.  
EXECUTE.
```

```
FILTER OFF.  
USE ALL.  
SELECT IF(g_final >-5).  
EXECUTE .
```

```
compute trial=studid.  
compute g=g_final.  
execute.
```

```
*DO NOT USE WITHOUT HELP***.  
matrix.  
get trial /var=trial.  
get g /var=g.  
get n1 /var=n1.  
get n2 /var=n2.
```

```

compute n=n1+n2.
compute k=nrow(n).

compute totaln=msum(n).

compute d=g*(1-(3/(4*(n)-9))).
compute vard=(n/(n1*n2))+d**2/(2*(n-3.94)).
compute wid=1/vard.
compute percwid=100*wid/msum(wid).
compute sed=sqrt(vard).
compute cilow=d-1.96*sed.
compute ciup=d+1.96*sed.
compute weight_d=msum(d*wid)/msum(wid).
compute se_d=sqrt(1/msum(wid)).
compute low_d=weight_d-1.96*se_d.
compute hig_d=weight_d+1.96*se_d.
compute ad_chi=(weight_d**2)*msum(wid).
compute ad_sig=1-chicdf(ad_chi,1).
compute hetd_chi=msum(wid*(d-weight_d)**2).
compute hetd_sig=1-chicdf(hetd_chi,k-1).
compute ofsn=k*(abs(weight_d/0.2)-1).
print k
/format="f8.0"
/title="Number of trials (k)".
print totaln
/format="f8.0"
/title="Total sample size (N)".
print /title="Standardized mean differences (Hedges unbiased g):".
print {d,sed,percwid,cilow,ciup}
/format="f10.4"
/title="Processed data"
/clabels="Hedges g" "se(g)" "%Weight" "95%Lower" "95%Upper".
print weight_d
/format="f8.3"
/title="Pooled Standardised Mean Difference".
print se_d
/format="f8.3"
/title="Standard error of estimator".
print {low_d,hig_d}
/format="f8.3"
/title="95% CI"
/clabels="Lower" "Upper".
print ad_chi
/format="f8.4"
/title="Association Chi-square statistic (df=1)".
print ad_sig
/format="f8.3"
/title="Significance of Chi-square (H0: No association)".
print ofsn
/format "f8.0"
/Title="Orwin's Fail Safe N for Cohen's small effect size (0.2)".
print hetd_chi
/format="f8.4"
/title="Chi-square of heterogeneity statistic (df=k-1)".

```

```

print hetd_sig
/format="f8.3"
/title="Significance of Chi-square (H0: Homogeneity)".
do if hetd_chi>k-1.
do if hetd_sig>0.10.
print /title="WARNING: Q p>0.10, but some heterogeneity exists".
end if.
compute h=sqrt(hetd_chi/(k-1)).
print h
/format="f8.2"
/title="H statistic (approx. ratio of CI lengths from random vs fixed-
+ "effects models)".
end if.
compute trial={trial;k+1}.
compute percwid={percwid;100}.
compute d={d;weight_d}.
compute sed={sed;se_d}.
compute procdata={trial,percwid,d,sed}.
compute vecnam={"trial","weights","d","se"}.
save procdata /outfile=* /names=vecnam.
end matrix.

* Thanks to Ray Levesque for his help at this step *.
FORMATS trial(F8.0).
compute nobreak=1.
match files file=* /by nobreak /last=last.
do if last=1.
write outfile='c:\temp\syntax.sps' /VALUE LABEL trial 'trial' "Total".'.
end if.
execute.
include 'c:\temp\syntax.sps'.

* Forest plot with individual and aggregated effect sizes *.
compute lowd=d-1.96*se.
compute highd=d+1.96*se.
VAR LABEL lowd 'Lower 95%CI' /highd 'Upper 95%CI' /d 'Effect size (g)'.
GRAPH /HILO(SIMPLE)=VALUE( highd lowd d ) BY trial.

```

APPENDIX C

BIBLIOGRAPHY OF STUDIES FOR META-ANALYSIS

- Al-Hadlaq, Mohammed (2003). Retention of words learned incidentally by Saudi EFL learners through working on vocabulary learning tasks constructed to activate varying depths of processing (Doctoral dissertation, Ball State University, 2003). *Dissertation Abstracts International*, 64. 05A.
- Ali Faqeehi, Muhammad. (2003). Exploring the effects of combined strategies on English vocabulary learning among Saudi Arabian University freshmen ( Doctoral dissertation, Ball State University, 2001). *Dissertation Abstracts International*, 64. 05A.
- Al-Seghayer, K. (2001). The effect of multimedia annotation modes on Ls vocabulary acquisition: A comparative study. *Language, Learning & Technology*, 5(1), 202-229.
- Atay, D. & Ozbulgan, C. (2006) Memory strategy instruction, contextual learning and ESP vocabulary recall. *English for Specific Purposes*, 26(1), 39-51.
- Berley, H. E. (1993). A meta-analysis of college developmental studies programs (remediation). Doctoral dissertation, Texas A&M University, 1993). *Dissertation Abstracts International*, 54, 05A.
- Bird, S. A. & Jacobs, G. M. (1999). An examination of the keyword method: how effective is it for native speakers of Chinese learning English? *Asian Journal of English Language Teaching*, 9, 75-97.
- Crow, J. T. & Quigley, J. R. (1985). A semantic field approach to passive vocabulary acquisition for reading comprehension. *TESOL Quarterly*, 19(3), 497-513.
- Ellis, R. & He, X. (1999). The role of modified input and output in the incidental acquisition of word meanings. *Studies in Second Language Acquisition*. 21. 258-301.
- Garcia Lopez, M. (2001). Helping secondary readers: the teacher as a living dictionary. *CAUCE*, 24, 679-691.
- Hill, M. & Laufer, B. (2003). Type of task, time-on-task and electronic dictionaries in incidental vocabulary acquisition. *IRAL*, 41, 87-106.
- Hulstijn, J & Laufer, B (2001) Some Empirical Evidence for the involvement load hypothesis in Vocabulary acquisition, *Language Learning*, 51(3 ), 539-558.
- Joe, A. (1998). What effects do text-based tasks promoting generation have on incidental vocabulary acquisition? *Applied Linguistics*, 19(3), 357-377.

- Kang, S. (1992). Computer and context-embedded language learning. *CAELL journal*, 3(3), 15-20.
- Kang, S. (1994). The effects of computer-based, context-embedded approaches to second-language vocabulary learning (Doctoral dissertation, University of Illinois at Urbana-Champaign, 1994). *Dissertation Abstracts International*, 55, 09A.
- Kim, Dae-sang (2006). Effects of text, audio, and graphic aids in multimedia instruction on the achievement of students in vocabulary learning (Doctoral dissertation, Indiana State University, 2006). *Dissertation Abstracts International*, 68, 02A.
- Laufer, B. (2003). Vocabulary acquisition in a second language: do learners really acquire most vocabulary by reading? Some empirical evidence. *The Canadian Modern Language Review*. 59(4), 567-587.
- Laufer, B. (2006). Comparing focus on form and focus on forms. *The Canadian Modern Language Review*. 63(10), 149-166.
- Laufer, B. (2005). Focus on form in second language vocabulary learning, *EUROSLA Yearbook*, 5. 223-250.
- Liu, Guiqiu (2006). Exploring the influence of western picture books on Chinese elementary students' vocabulary gains, story-retelling abilities, and attitudes toward English Language learning (Doctoral dissertation, Indiana University of Pennsylvania, 2006). *Dissertation Abstracts International*, 67, 02A.
- Luppescu, S. & Day, R.R.(1993). Reading, Dictionaries, and Vocabulary learning. *Language Learning*, 43(2), 263-287.
- Oskarsson, M. (1975). On the role of the mother tongue in learning foreign language vocabulary: an empirical investigation. *ITL, Review of Applied Linguistics*, 27, 19-32.
- Qian, D.D.(1996). ESL Vocabulary Acquisition: Contextualization and decontextualization. *Canadian Modern Language Review*, 53, 120-142.
- Rodriguez, M & Sadoski, M. (2000). Effects of rote, context, keyword, and context/keyword methods on retention of vocabulary in EFL classroom. *Language Learning*, 50, 385-412.
- Smith, B. (2004). Computer-mediated negotiated interaction and lexical acquisition. *Studies in Second Language Acquisition*, 26, 365-398.



- Srinaowaratt, Sirinan (1991). The effects of two vocabulary instruction programs on vocabulary learning of eleventh grade Thai EFL students (Doctoral dissertation, Illinois State University, 1991). *Dissertation Abstracts International*, 53, 02A.
- Sun, Y. & Dong, Q. (2004). An Experiment on supporting children's English vocabulary learning in multimedia context. *Computer Assisted Language Learning*, 17(2), 131-147.
- Tozcu, A. & Coady, J. (2004). Successful learning of frequent vocabulary through CALL also benefits reading comprehension and speed. *Computer Assisted Language Learning*, 17(5), 473-495.
- Van Bussel, F.J. (1994). Design rules for computer-aided learning of vocabulary items in a second language. *Computers in Human Behavior*, 10, 63-76.
- Vespor, M. & Lowie, W. (2003). Making sense of polysemous words. *Language Learning*, 53(3), 547-586.
- Wang, Xigu. (2001). Standard definitions versus revised definitions: Their effectiveness in understanding word meaning with Chinese adult learners of English as a foreign language (Doctoral dissertation, Oakland University, 2001). *Dissertation Abstracts International*, 62, 07A.
- Yeung, A. S. (1999). Cognitive load and learner expertise: split-attention and redundancy effects in reading comprehension tasks with vocabulary definitions, *Journal of Experimental Education*, 67(3), 197-218.

APPENDIX D 1

CLASSIFICATION OF STUDIES

Contextualized Instruction

<i>Author(s)</i>	<i>N</i>	<i>ES</i>
Laufer (2003)	72	1.93
Rodriguez&Sadoski (2000)	35	-0.27
Sun & Dong (2004)	34	-0.25
Ali Faqeehi (2003)	62	0.26
Kim (2006)	65	-0.44
Joe (1998)	32	1.84
Hulstijn & Laufer (2001)	54	1.25
Al-Hadlaq (2003)	55	0.67
Kang (1994)	48	0.77

K=9, ES=.58

\* N= number of samples  
 ES=effect size  
 K= number of studies

APPENDIX D 2

CLASSIFICATION OF STUDIES

Semi-contextualized Instruction

<i>Authors</i>	<i>N</i>	<i>ES</i>
Ellis&He (1999 )	34	.67
Laufer (2003)	52	1.34
Laufer (2003)	60	2.31
Al-Seghayer (2001)	60	1.07
Smith (2004)	36	.69
Laufer (2005)	52	1.35
Laufer (2005)	96	.55
Ali Faqeehi (2003)	63	-.09
Kim (2006)	77	.37
Crow&Quigley (1985 )	42	-.40
Bird & Jacobs (1999)	49	.20
Atay & Ozbulgan (2006)	50	.95
Kang (1992)	30	.48
Van Bussel (1994)	29	.99
Hulstijn & Laufer (2001)	53	.65
Wang (2001)	96	.08
Wang (2001)	96	.84
Liu (2006)	72	1.97

Srinaowaratt (1991)	52	.45
Kang (1994)	46	.06

K=20, ES=.65

\* N= number of samples  
ES=effect size  
K= number of studies

APPENDIX D 3

CLASSIFICATION OF STUDIES

Decontextualized Instruction

<i>Author(s)</i>	<i>N</i>	<i>ES</i>
Hill & Laufer (2003)	65	0.55
Sun & Dong (2004)	32	2.25
Vespor&Lowie (2003)	78	0.59
Quin (1996)	63	0.77
Laufer (2005)	158	1.11
Yeung (1999)	96	0.42
Yeung (1999)	84	1.71
Laufer (2006)	158	1.00
Lupescu & Day (1993)	297	0.77
Oskarsson (1975)	184	0.47
Garcia Lopez (2001)	40	-0.2
Kang (1994)	48	0.25

K=12, ES=.76

- \* N= number of samples
- ES=effect size
- K= number of studies

APPENDIX D 4

CLASSIFICATION OF STUDIES

ESL condition

Author(s)	N	ES
Ellis&He (1999)	34	0.67
Al-Seghayer (2001)	60	1.07
Smith (2004)	36	0.69
Crow&Quigley (1985)	42	-0.4
Kang, Sook-hi (1992)	30	0.48
Joe (1998)	32	1.84

K=6, ES = .66

\* N= number of samples  
ES=effect size  
K= number of studies

APPENDIX D 5

CLASSIFICATION OF STUDIES

EFL condition

Author(s)	N	ES
Laufer (2003)	52	1.34
Laufer (2003)	72	1.93
Laufer (2003)	60	2.31
Hill & Laufer (2003)	65	0.55
Rodriguez&Sadoski (2000)	35	-0.27
Sun & Dong (2004)	34	-0.25
Sun & Dong (2004)	32	2.25
Vespor&Lowie (2003)	78	0.59
Quin (1996)	63	0.77
Laufer (2005)	52	1.35
Laufer (2005)	96	0.55
Laufer (2005)	158	1.11
Ali Faqeehi (2003)	63	-0.09
Ali Faqeehi (2003)	62	0.26
Kim (2006)	65	-0.44
Kim (2006)	77	0.37
Yeung (1999)	96	0.42
Yeung (1999)	84	1.71

Laufer (2006)	158	1
Bird & Jacobs (1999)	49	0.2
Atay & Ozbulgan (2006)	50	0.95
Luppescu & Day (1993)	297	0.77
Oskarsson (1975)	184	0.47
Garcia Lopez (2001)	40	-0.2
Van Bussel (1994)	29	0.99
Hulstijn & Laufer (2001)	53	0.65
Hulstijn & Laufer (2001)	54	1.25
Wang (2001)	96	0.08
Wang (2001)	96	0.84
Liu (2006)	72	1.97
Al-Hadlaq (2003)	52	0.67
Srinaowaratt (1991)	52	0.45
Kang (1994)	48	0.25
Kang (1994)	46	0.06
Kang (1994)	48	0.77

K=35, ES = .69

\* N= number of samples

ES=effect size

K= number of studies



APPENDIX D 6

CLASSIFICATION OF STUDIES

Studies with Elementary Students

Author	N	ES
Yeung (1999)	96	0.42
Liu (2006)	72	1.97
Kang (1994)	48	0.25
Kang (1994)	46	0.06
Kang (1994)	48	0.77

K=5, ES=.65

\* N= number of samples  
ES=effect size  
K= number of studies

APPENDIX D 7

CLASSIFICATION OF STUDIES

Studies with Secondary Students

Authors	N	ES
Laufer (2003)	60	2.31
Rodriguez&Sadoski (2000)	35	-0.27
Laufer (2005)	158	1.11
Kim (2006)	65	-0.44
Kim (2006)	77	0.37
Yeung (1999)	84	1.71
Laufer (2006)	158	1
Garcia Lopez (2001)	40	-0.2
Van Bussel (1994)	29	0.99
Srinaowaratt (1991)	52	0.45

K=10, ES= .78

\* N= number of samples

ES=effect size

K= number of studies

APPENDIX D 8

CLASSIFICATION OF STUDIES

Studies with College Students

Authors	N	ES
Ellis&He (1999)	34	0.67
Laufer (2003)	52	1.34
Laufer (2003)	72	1.93
Hill & Laufer (2003)	65	0.55
Al-Seghayer (2001)	60	1.07
Vespor&Lowie (2003)	78	0.59
Quin (1996)	63	0.77
Smith (2004)	36	0.69
Laufer (2005)	52	1.35
Laufer (2005)	96	0.55
Ali Faqeehi (2003)	63	-0.09
Ali Faqeehi (2003)	62	0.26
Crow&Quigley (1985)	42	-0.4
Bird & Jacobs (1999)	49	0.2
Atay & Ozbulgan (2006)	50	0.95
Luppescu & Day (1993)	297	0.77
Joe (1998)	32	1.84

Oskarsson (1975)	184	0.47
Hulstijn & Laufer (2001)	53	0.65
Hulstijn & Laufer (2001)	54	1.25
Wang (2001)	96	0.08
Wang (2001)	96	0.84
Al-Hadlaq (2003)	52	0.67

K=23, ES=.67

\* N= number of samples  
 ES=effect size  
 K= number of studies

APPENDIX D 9

CLASSIFICATION OF STUDIES

Advanced English Learners

Author(s)	N	ES
Laufer (2003)	52	1.34
Laufer (2003)	72	1.93
Hill & Laufer (2003)	65	0.55
Crow&Quigley (1985)	42	-0.4
Hulstijn & Laufer (2001)	53	0.65
Hulstijn & Laufer (2001)	54	1.25

K=6, ES=.88

\* N= number of samples  
ES=effect size  
K= number of studies

APPENDIX D 10

CLASSIFICATION OF STUDIES

Intermediate English Learners

Author(s)	N	ES
Ellis&He (1999)	34	0.67
Laufer (2003)	60	2.31
Al-Seghayer (2001)	60	1.07
Smith (2004)	36	0.69
Ali Faqeehi (2003)	63	-0.09
Ali Faqeehi (2003)	62	0.26
Bird & Jacobs (1999)	49	0.2
Joe (1998)	32	1.84
Wang (2001)	96	0.08
Wang (2001)	96	0.84
Al-Hadlaq (2003)	52	0.67

K=11, ES=.61

\* N= number of samples  
 ES=effect size  
 K= number of studies

APPENDIX D 11

CLASSIFICATION OF STUDIES

Multimedia Use

Author(s)	N	ES
Hill & Laufer (2003)	65	0.55
Al-Seghayer (2001)	60	1.07
Sun & Dong (2004)	34	-0.25
Sun & Dong (2004)	32	2.25
Smith (2004)	36	0.69
Laufer (2005)	96	0.55
Kim (2006)	65	-0.44
Kim (2006)	77	0.37
Tozcu & Coady (2004)	56	3.42
Kang (1994)	48	0.25
Kang (1994)	46	0.06
Kang (1994)	48	0.77
Kang (1992)	30	0.48

K=13, ES=.61

\* N= number of samples  
 ES=effect size  
 K= number of studies

APPENDIX D 12

CLASSIFICATION OF STUDIES

First Language Supports

Author(s)	N	ES
Laufer (2003)	72	1.93
Laufer (2003)	60	2.31
Hill & Laufer (2003)	65	0.55
Rodriguez&Sadoski (2000)	35	-0.27
Sun & Dong (2004)	32	2.25
Vespor&Lowie (2003)	78	0.59
Laufer (2005)	52	1.35
Laufer (2005)	96	0.55
Laufer (2005)	158	1.11
Ali Faqeehi (2003)	63	-0.09
Ali Faqeehi (2003)	62	0.26
Laufer (2006)	158	1
Bird & Jacobs (1999)	49	0.2
Atay & Ozbulgan (2006)	50	0.95
Luppescu & Day (1993)	297	0.77
Oskarsson (1975)	184	0.47
Garcia Lopez (2001)	40	-0.2
Van Bussel (1994)	29	0.99



---

Hulstijn & Laufer (2001)	53	0.65
Hulstijn & Laufer (2001)	54	1.25
Wang (2001)	96	0.08
Kang (1994)	48	0.25
Kang (1994)	46	0.06
Kang (1994)	48	0.77

---

K=24, ES=.69

\* N= number of samples  
ES=effect size  
K= number of studies