



Theses and Dissertations

2008-07-11

The Relationship Between Vocabulary Knowledge and Reading Comprehension of Authentic Arabic Texts

Shereen Maher Salah
Brigham Young University - Provo

Follow this and additional works at: <https://scholarsarchive.byu.edu/etd>



Part of the [Other Languages, Societies, and Cultures Commons](#)

BYU ScholarsArchive Citation

Salah, Shereen Maher, "The Relationship Between Vocabulary Knowledge and Reading Comprehension of Authentic Arabic Texts" (2008). *Theses and Dissertations*. 1827.
<https://scholarsarchive.byu.edu/etd/1827>

This Thesis is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of BYU ScholarsArchive. For more information, please contact scholarsarchive@byu.edu, ellen_amatangelo@byu.edu.

THE RELATIONSHIP BETWEEN VOCABULARY KNOWLEDGE AND READING
COMPREHENSION OF AUTHENTIC ARABIC TEXTS

By

Shereen Maher Salah

A thesis submitted to the faculty of

Brigham Young University

in partial fulfillment of the requirements for the degree of

Master of Arts

Center for Language Studies

Brigham Young University

August 2008

Copyright © 2008 Shereen Maher Salah

All Rights Reserved

BRIGHAM YOUNG UNIVERSITY

GRADUATE COMMITTEE APPROVAL

of a thesis submitted by

Shereen Maher Salah

This thesis has been read by each member of the following graduate committee and by majority vote has been found to be satisfactory.

Date

R. Kirk Belnap, Chair

Date

Dilworth B. Parkinson

Date

J. Paul Warnick

BRIGHAM YOUNG UNIVERSITY

As chair of the candidate's graduate committee, I have read the thesis of Shereen Maher Salah in its final form and have found that (1) its format, citations, and bibliographical style are consistent and acceptable and fulfill university and department style requirements; (2) its illustrative materials including figures, tables, and charts are in place; and (3) the final manuscript is satisfactory to the graduate committee and is ready for submission to the university library.

Date

R. Kirk Belnap
Chair, Graduate Committee

Accepted for the Department

Ray T. Clifford
Director, Center for Language Studies

Accepted for the College

Joseph Parry
Associate Dean, College of Humanities

ABSTRACT

THE RELATIONSHIP BETWEEN VOCABULARY KNOWLEDGE AND
READING COMPREHENSION OF AUTHENTIC ARABIC TEXTS

Shereen Maher Salah

Center for Language Studies

Master of Arts

This study investigates the relationship between vocabulary knowledge and reading comprehension of authentic Arabic texts; in particular, it attempts to investigate the percentage of vocabulary coverage (known words) readers need to ensure reading comprehension of two reading passages from online Arabic news source. Data was collected from twenty-three Arabic as-foreign language (AFL) learners at Brigham Young University, who ranged from Intermediate Low to Intermediate Mid in both productive and receptive skills. Two reading comprehension tests, circling the unknown words in texts and a lexical coverage test for each passage texts were given to the subjects.

A linear regression analysis of the data shows that there is a correlation coefficient of 0.7 and 0.6 between the percentage of known words and

students' comprehension of the two reading texts. The results indicate that the subjects needed to know approximately 90% of running words to adequately comprehend the first passage and around 86% to comprehend the second passage. Based on the findings, this study suggests that there is a lexical threshold for AFL learners, below which adequate comprehension of authentic texts might not be possible.

DEDICATION

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

اقْرَأْ بِاسْمِ رَبِّكَ الَّذِي خَلَقَ ﴿١﴾ خَلَقَ الْإِنْسَانَ مِنْ عَلَقٍ ﴿٢﴾ اقْرَأْ وَرَبُّكَ

الْأَكْرَمُ ﴿٣﴾ الَّذِي عَلَّمَ بِالْقَلَمِ ﴿٤﴾

Read in the name of your Lord Who created. (1) He created man from a clot. (2) Read and your Lord is Most Honorable, (3) Who taught (to write) with the pen (4)

To my Father.

ACKNOWLEDGMENTS

As always, first: Thank God for giving me the chance for better education; i.e. one of numerous privileges I am blessed with.

All I am, I owe to you, my mother, I would not achieve anything in this life, without you, your unconditional support and ceaseless prayers. You are the best.

Ahmad, my friend, my love, my life-long partner and my in-house statistical consulting center: I could not finish this thesis, without your valuable help. You almost wrote this acknowledgment. Thank you for being so patient, understanding and supportive. I LOVE YOU.

Wholehearted thanks go to my “singing stars”; Sama and Noureen. You were not too young to help me. Thanks for adding more joy in my life, thanks for your smiles and challenges.

I am indebted to my thesis advisor, Kirk Belnap for his continuous support and encouragement. Many thanks go to my committee members, Dilworth Parkinson and Paul Warnick, for their valuable comments and suggestions.

TABLE OF CONTENTS

Chapter 1: Introduction.....	1
Vocabulary Knowledge	3
Vocabulary Knowledge and Reading Comprehension	3
Authenticity of L2 Reading Materials	5
Rationale for this Study	8
Research Questions.....	8
Significance of the Study for AFL Teachers	9
Significance of the Study for AFL Research	9
Chapter 2: Literature Review	10
L2 Reading Models.....	10
Effects of Vocabulary Knowledge on L2 Reading Comprehension.....	13
Definition of Knowing a Word.....	13
Vocabulary Size and Reading Comprehension	16
Vocabulary Knowledge in AFL.....	17
Vocabulary Knowledge and Reading Comprehension in AFL	19
L2 Linguistic Threshold.....	20
The Definition of Linguistic Threshold.....	20
The Development of the Notion of Linguistic Threshold	21
Lexical Threshold for L2 Reading Comprehension	21
Authenticity in L2 Reading.....	24

Definition of Authentic Materials.....	24
Comprehensibility of Authentic Materials in L2.....	26
Chapter 3: Methodology.....	30
The Subjects.....	31
Material.....	32
Instruments.....	33
Background Questionnaire	33
Reading Comprehension Test (RCT)	33
Identifying Unknown Vocabulary	34
Lexical Coverage Test (LCT).....	35
Data Collection Procedures.....	36
Session One	36
Session Two.....	37
Research Design.....	37
Null Hypotheses.....	37
Variables	38
Dependent Variable	38
Independent Variable.....	38
Statistical Analysis.....	39
Chapter 4: Results and Discussion	41
Mean and Standard Deviation of the Dependent Variable	42
Mean and Standard Deviation of the Independent Variable.....	44

Validity of the Statistical Procedures.....	45
Variables Correlation	47
Passage One	47
Passage Two	50
Density and Coverage	54
Chapter 5: Conclusions.....	57
Limitation of the Study	59
Implications of the Study	60
Implications for AFL Instructors and Teachers:.....	60
Recommendations for Future Research	61
Bibliography	63
Appendix A: Background Questionnaire	72
Appendix B: Reading Comprehension Test - Passage 1.....	73
Appendix C: Reading Comprehension Test – Passage 2.	76
Appendix D: Lexical Coverage Test – Passage 1.	79
Appendix E: Lexical Coverage Test - Passage 2.....	81

LIST OF FIGURES

Figure 1: RCT Scores – Passage 1.	43
Figure 2: RCT Scores – Passage 2.	43
Figure 3: Distribution of Known Words Percentage among Subjects - Passage 1	45
Figure 4: Distribution Known Words Percentage among Subjects - Passage 2.....	45
Figure 5: Normal Probability Plot – Passage 1.	46
Figure 6: Normal Probability Plot – Passage 2.	46
Figure 7: Percentage Known Words vs. Comprehension Score – Passage 1.....	48
Figure 8: Percentage Known Words vs. Comprehension Score – Passage 2.....	51

LIST OF TABLES

Table 1 : Age Range of Subjects.	32
Table 2: L2 Background Knowledge (Excluding Arabic).....	32
Table 3: Mean & SD of Reading Comprehension Scores.....	42
Table 4: Mean & SD of Word Coverage Percentage.	44
Table 5: Regression Analysis Results – Passage 1.....	48
Table 6: Minimum Percentage of Known Words – Passage 1.....	49
Table 7: Regression Analysis Results – Passage 2.....	50
Table 8: Minimum Percentage of Known Words – Passage 2.....	52
Table 9: Minimum Percentage of Known Words – Mean, Both Passages.	53
Table 10: Density of Unknown Words to Total Words.	54

Chapter 1: Introduction

Reading is one of the five language skills (listening, writing, reading, speaking, and culture) that second language learners (SLL) should acquire in their language learning process if they are to become well-rounded users of the target language. The ability to read is seen as the most stable and durable of the second language skills (Rivers, 1981). Language learners acquire most of their vocabulary through reading, particularly if they do not stay in a country where that language is spoken. SLL can lose their writing and speaking skills, but still be able to comprehend text with some degree of proficiency (Bernhardt, 1991).

Reading helps in improving cultural skills. Reading authentic materials provides SLL with good sources of the target language culture through reading about different topics and aspects of the culture. Also reading improves the productive skill of writing. It can help learners to improve their writing skills as sometimes they are subconsciously acquire good writing style from reading authentic written texts which are both “highly accessible and cheap sources of second language materials” (Bernhardt, 1991, p. 1). Through reading L2 learners have the opportunity to be exposed to well-organized and well-written pieces of writing which help them to improve their language abilities and to build writing schemata (Duran, 1999). “Reading may be regarded as a necessary precondition

for any writing task, since the writer must be aware of the structure of a particular type of writing before he can produce it” (Kennedy & Bolitho, 1985, p. 85).

Like readers who are native speakers, when SLLs read a text their main goal is to comprehend its meaning. Traditionally reading was seen as a receptive skill and the readers are passive recipients, but, in fact, reading is a process of interaction between the text and the reader. “The reader interacts with the text to create meaning as the reader’s mental processes interact with each other at different levels (e.g. letter, lexical, syntactic, or semantic) to make the text meaningful” (Barnett, 1989, p. 29). The reader actually involved in an active and constructive process, building meaning from a text. “Meaning does not exist in a text but in readers and the representations they build” (Hass & Flower, 1988, p. 167).

Reading components are text-based (vocabulary, syntax, rhetorical structure and cultural content) and reader-based (prior knowledge, cognitive development, interest and purpose in reading, and reading strategies) (Barnett, 1989). Researchers have suggested that among the text-based components, vocabulary is the most important factor in reading comprehension (Laufer, 1989; Laufer & Sim, 1985; Nation, 1990).

This thesis investigates the relationship between vocabulary knowledge and reading comprehension in Arabic as a foreign language (AFL). Specifically, the study assesses the percentage of known words needed for intermediate AFL learners to gain adequate comprehension of authentic Arabic texts (news articles).

Vocabulary Knowledge

While there is agreement on the importance of vocabulary knowledge in reading comprehension (Alderson, 2000; Anderson & Freebody, 1981), researchers disagree about what it means to fully know a word and what kind of knowledge this is. Cronbach (1942) categorizes vocabulary as the knowledge of word meaning and the level of one's accessibility to this knowledge, but this definition ignores other aspects of lexical knowledge such as spelling, pronunciation, and morpho-syntactic properties (as cited in Qian 2002). Richards (1976) offered the first inclusive definition of vocabulary knowledge, which not only included the morphological and syntactic properties but also other aspects, such as word frequency. Yet his definition was still missing the pronunciation and spelling aspects. In 1990, Nation, however, included these missing aspects in his framework of vocabulary knowledge. He argued that a person's knowledge of a word should entail both receptive and productive knowledge, all aspects of what is involved in knowing a word, which includes forms, meaning and usage.

Vocabulary Knowledge and Reading Comprehension

Vocabulary knowledge and its role in reading comprehension has been one of the main areas of focus in second language research for the last twenty years. Both vocabulary knowledge and reading comprehension are closely related, and this relationship is not one-directional, since vocabulary knowledge can help

the learner to comprehend written texts and reading can contribute to vocabulary growth (Chall, 1987; Nation, 2001; Stahl, 1990).

Some researchers advocate that vocabulary is the most crucial factor in reading comprehension. Cooper (1984) described vocabulary as being the key ingredient to successful reading while other researchers argue that “no text comprehension is possible, either in one’s native language or in a foreign language, without understanding the text’s vocabulary” (Laufer 1997, p. 20). They maintain that when the percentage of unknown vocabulary in a given text increases, the possibility of comprehending the text decreases (Hirsh & Nation, 1992; Hu & Nation, 2000; Laufer, 1989, 1992, 1997).

According to her research, Laufer (1989) was more specific when she revealed the importance of having sufficient vocabulary for reading comprehension, claiming that a reader whose vocabulary is insufficient to cover at least 95% of the words in a passage will not be guaranteed comprehension. Readers themselves consider vocabulary knowledge to be the main obstacle to second language reading comprehension. Yorio (1971) surveyed second language students, who reported that vocabulary was their main problem in reading comprehension.

There is on-going disagreement over whether there is a threshold of vocabulary knowledge below which appropriate reading comprehension is not guaranteed for L2 learners. Nation (2001) notes that the linguistic threshold can be viewed in two ways: The traditional meaning is that “If a learner has not crossed the threshold, then adequate comprehension is not possible,” and the other

meaning considers the threshold “as a probabilistic boundary. That is, if a learner has not crossed the threshold, the chances of comprehending adequately are low” (p. 144).

Authenticity of L2 Reading Materials

The idea of authenticity and its importance in language learning has a long history. The current push for authenticity dates back to 1970, when the communicative language teaching (CLT) approach was first introduced, rejecting the strictly structured approaches of language learning that had been the rule (Mishan, 2005). CLT assumes that there is an underlying ability for language use that involves far more than knowledge of the grammar of the language (Brumfit & Johnson, 1979). “There are rules of use without which the rules of grammar would be useless” (Hymes, 1971, p. 15). These rules of use are part of the communicative competence that L2 learners need to know about the language and its culture and how to use the language to communicate successfully. Those who employed the CLT approach and presented the authentic materials to L2 learners realized the large gap between the text-books that only present artificial language and the real language that authentic texts provide. Swaffar (1985) argued that some foreign language textbooks do not reflect any communicative goal and that some reading materials are missing “the linguistically authentic comprehensible input presented in a fashion which allows students to practice decoding message systems rather than individual words” (p. 17).

Despite the difficulties associated with authentic texts, like the ambiguity of some cultural concepts and the typical oversupply of new vocabulary, many advantages have been noticed. In 1981, Swaffar argued that “the sooner the students are exposed to authentic language, the more rapidly they will learn that comprehension is not a function of understanding every word, but rather of developing strategies..., strategies essential in both oral and written communication” (p. 188). Unlike simplified or edited materials, which typically do not focus on the actual reading–comprehension process (Bernhardt, 1984), some studies have shown that using authentic texts has a positive effect on learning the target language by developing communicative competence (Peacock, 1997). Unlike authentic materials, simplified or edited materials are made expressly to limit the L2 learners to vocabulary and morphological structures that they know or that the writer wants them to encounter (Laufer, 1989).

Mishan (2005), in her book, *Designing Authenticity into Language Learning*, developed the notion of the 3Cs, culture, currency and challenge that summarizes the advantages of authentic texts over those written for L2 learners. Mishan feels that authentic materials are a representation of the target language culture. “Culture and language are indivisible, any and every linguistic product of a society from a newspaper headline to a food label embody/represent the culture” (p. 44). The currency of authentic materials is a unique feature, especially those from the media, as they “offer topics and language in current use, as well as those relevant to the learners” (p. 44). Also authentic texts can be seen as a positive

challenge to learners. “Authentic texts are intrinsically more challenging yet can be used at all proficiency levels” (p. 44).

Authentic materials are also seen by some researchers as a motivating force. Most students learn a language to communicate in the real language and this is the opportunity that authentic texts provide to the learner. It is remarkable how students get excited when they are able to comprehend a reading passage from an Arabic newspaper and add a new word to their vocabulary knowledge. They feel that it is a step closer towards proficiency. In addition, Little (1997) argues that exposure to and familiarity with authentic texts can also help to instill confidence in the face of the target language (p. 231), which is what most L2 learners are looking for in the L2 acquisition process.

Belnap (2006) investigated learners’ motivations and goals for learning Arabic and found that most of the students learn Arabic to read the modern Arabic press and to interact with Arabs. In order for learners to achieve their learning goal, they need a good diet of real language. Sweet (1899) pointed out that “the great advantage of natural, idiomatic texts over artificial ‘methods’ or ‘series’ is that they do justice to every feature of the language. The artificial systems, on the other hand, tend to cause incessant repetition of certain grammatical constructions, certain elements of the vocabulary, certain combinations of words to the almost total exclusion of others which are equally, or perhaps even more, essential” (p. 177).

Rationale for this Study

There is agreement among second language researchers that vocabulary knowledge is an important component in reading comprehension. As noted, many studies of English as a foreign language have suggested that L2 readers must understand 95% of the words in any text to ensure reasonable reading comprehension of the text (Laufer, 1989; Laufer & Sim, 1985). Other researchers have suggested that L2 learners need 98% of word coverage to read un-simplified texts for pleasure (Hirsh & Nation, 1992; Nation, 2000). Little research has been conducted in this area regarding Arabic as a second language. Few studies have investigated the relationship between vocabulary and reading comprehension and they have not investigated the percentage of word coverage needed to assure reading comprehension nor the possibility of a lexical threshold below which the L2 learner cannot achieve the minimum level of reading comprehension.

Therefore, there is a need for empirical studies investigating the percentage of vocabulary coverage needed for AFL learners to ensure reading comprehension of authentic Arabic texts (news articles).

Research Questions

This study is one step in this direction. It attempts to answer the following research questions:

- Is there a correlation between the percentage of known vocabulary and reading comprehension of authentic Arabic texts?
- Is there a minimum percentage of known vocabulary that acts as a lexical threshold for AFL (ranging between Intermediate-Low to Intermediate-Mid) to effectively read Arabic news articles?

Significance of the Study for AFL Teachers

This research could demonstrate to Arabic teachers the importance of teaching vocabulary. It could also help teachers in selecting authentic reading passages that are appropriate for certain levels of learners based on the percentage of known/unknown vocabulary. It will also draw teachers' attention to the necessity of teaching reading strategies such as the guessing strategy to recognize unknown vocabulary. It also emphasizes that second language readers need to use contextual information to guess/understand unknown vocabulary.

Significance of the Study for AFL Research

This study provides new results on the relationship between vocabulary knowledge and comprehension of authentic Arabic texts. While research has been done on the subject, studies directed specifically at Arabic are scarce. This study is expected to enrich AFL reading research and encourage other researchers of Arabic to further investigate this important topic.

Chapter 2: Literature Review

The main purpose of this study is to investigate the relationship between vocabulary knowledge and reading comprehension in Arabic as a foreign language. Specifically, it evaluates the effect of unknown vocabulary on intermediate-level AFL learners' ability to comprehend authentic Arabic texts. In other words, it will attempt to investigate the percentage of vocabulary coverage readers need to ensure comprehension.

It should be mentioned that this particular topic has not been covered in the currently available AFL literature. As a result, the literature review will focus on related literature from studies of other languages. This chapter explores literature on:

1. L2 reading models
2. The effects of vocabulary knowledge on L2 reading comprehension,
3. The nature of a linguistic threshold
4. Authenticity in L2 reading

L2 Reading Models

In the last 40 years, reading researchers have been studying the reading process. Based on their interpretation of the reading process they have developed

models for reading and defined the reading process in an attempt to depict how an individual perceives a word, processes a clause, and comprehends a text. (Singer & Ruddell, 1985).

Goodman (1968) explains the reading process as a “psycholinguistic guessing game” (p. 126) in which the reader uses “general knowledge of the world or of particular textual components to make intelligent guesses about what might come next in the text [and] samples only enough of the text to confirm or reject these guesses” (Barnett, 1989, p. 13).

Researchers tend to classify the models of reading into three categories:

1. Top-down
2. Bottom –up
3. Interactive

These reading models are used in first language (L1) reading, as well as in second or foreign language (SL/FL) reading (Barnett, 1989). The top-down model suggests that comprehension begins in the mind of the reader, who has already some ideas and prior knowledge about the meaning of the text. This model considers the readers and their interests, world knowledge, and reading skills as the driving force behind reading comprehension (Barnett, 1989; Goodman, 1968), as they construct meaning from their own previously acquired knowledge, as the text has no meaning in and of itself (Carrell & Eisterhold, 1983; Rumelhart, 1980).

According to the bottom-up model, comprehension begins with the processing of the smallest linguistic units, working towards larger units. This model focuses mainly on the written text, rather than the readers themselves.

In this model, the reading process is considered a text-driven, decoding process wherein the role of the reader is to reconstruct the meaning found in the text (Carrell, 1983; Gascoigne 2005; Gough, 1972). The text is a “chain of isolated words, each is deciphered individually” (Martinez-Lang, 1995, p. 70), and the reader is someone who “approaches the text by concentrating exclusively on the combination of letters and words in a purely linear manner” (p. 70).

While the top-down model focuses on reader variables (such as background knowledge, strategy use, reading purpose, and interest in the topic) as the main source of reading comprehension and the bottom-up model places an emphasis on text-based variables (vocabulary, syntax, grammatical structure...), the interactive model, the most recent of reading models, combines the top-down and bottom-up variables.

The interactive model is an interaction between top-down and bottom-up. This model suggests that reading comprehension is the result of interaction between the reader and the text (Bernhardt, 1991; Eskey, 1988; Grabe, 1991). Barnett (1989) describes this interaction as when “the reader interacts with the text to create meaning as the reader’s mental processes interact with each other at different levels to make the text meaningful” (Barnett 1989, p. 29). Many L2 reading researchers have accepted this interactive model, and several studies have been conducted to more closely examine the interactive model and the impact of

its (text-based and reader-based) variables on L2 reading comprehension. (Akagawa, 1995; Bossers 1991).

Effects of Vocabulary Knowledge on L2 Reading Comprehension

Definition of Knowing a Word

Many researchers have offered definitions for word comprehension. Grendel (1993) defines it as “knowing the meaning of a word” (p. 141). Vygotsky (1986) claims that “a word without meaning is an empty sound, no longer a part of human speech” (p. 6). As pointed out in Chapter One, Nation (2001) presents a word knowledge framework suggesting that a person’s knowledge of a word should be both receptive and productive, “to cover all aspects of what is involved in knowing a word” (Nation, 2001, p. 26). According to him, there are nine aspects of knowing a word that include form, meaning and use, as follows:

1. Spoken form
2. Written form
3. Word parts
4. Connection of form and meaning
5. Conceptual meaning
6. Association with related words
7. Grammatical functions
8. Collocation behaviors
9. Word usage constraints; appropriateness

Nation relates these nine aspects to the receptive and productive knowledge of words. Receptive knowledge is important for recognizing the meaning of a word without the need to produce the word again. On the other hand, productive knowledge is the knowledge that learners need to “produce language forms by speaking and writing to convey a message to others.” (p. 24).

Also, Ruddle (1994) divided knowing a word into six categories:

1. Knowing the word meaning aurally
2. Knowing the word meaning but not expressing it
3. Knowing the meaning but not the word
4. Knowing the partial meaning of the word
5. Knowing a different meaning of a word
6. Not knowing the concept or the label.

Thus, one’s knowledge of a word does not have to include both receptive and productive control to perform certain tasks. Schmitt & Meara (1997) argue that native speakers do not master all types of word knowledge. They only master a limited number of word knowledge categories for most of their lexicon, and have only the receptive knowledge of some low frequency words. Second language learners, like native speakers, do not have to have a full knowledge of all vocabulary to function in the language. This knowledge varies depending on the task the person is performing (Mazynski, 1983; Qian, 2002). Some activities involve only receptive knowledge while others require productive knowledge as well. In reading tasks, only receptive knowledge is required. L2 learners do not

need full knowledge of a word meaning for every reading comprehension task; partial knowledge may suffice. Through this partial or incomplete knowledge learners can gain additional knowledge about a word, as suggested by Henriksen (1999) when he proposed three components in vocabulary development:

1. Partial to precise knowledge
2. Depth of knowledge
3. Receptive to productive ability.

As indicated, L2 readers need to develop both receptive and productive vocabulary knowledge and to increase their vocabulary size. “When readers increase their vocabulary size, their use of language skills implicitly increases and their knowledge of the world also becomes broader” (Huang, 1999, p. 43). A larger vocabulary enhances other language skills of L2 learners. Language learners must acquire as much vocabulary as possible in order to effectively read in the language (Bernhardt & Kamil, 1995).

In 2001, Calfee, Graves, Ryder, & Slater listed five stages of vocabulary knowledge:

1. Learning to read a known word.
2. Learning new meanings of known words
3. Learning new words that represent known concepts
4. Clarifying and enriching meaning of known words
5. Moving words from receptive (listening and reading) to expressive (speaking and writing) vocabulary (p. 81).

The above definitions of vocabulary knowledge strengthen the argument that there is a relationship between vocabulary and reading comprehension.

Vocabulary Size and Reading Comprehension

As indicated by the researchers cited in the previous section, “reading a language” and comprehending it require that one possess sufficient vocabulary. The relation between vocabulary knowledge and reading comprehension is complex and dynamic. One way of looking at it is to divide it up into two major directions of effect. “The effect of vocabulary knowledge on reading comprehension, and the effect of reading comprehension on vocabulary knowledge or growth” (Nation & Hu, 2000, p. 403).

Vocabulary knowledge plays an important role in reading comprehension. Researchers tend to agree that vocabulary knowledge is a major prerequisite and causal factor in comprehension and that there is a relationship between vocabulary size and reading comprehension. Some studies have investigated this relationship and used vocabulary size as a predictor variable for reading comprehension (Hu & Nation, 2000; Laufer 1989, 1992; Liu Na & Nation, 1985).

Liu Na & Nation (1985) investigated this relationship by asking 59 ESL learners to recognize the meaning of the vocabulary in two texts with different vocabulary densities: 96% of the vocabulary in the first passage was known, whereas only 90% was in the second text. In their study, they found that the density of unknown words in the text affected the guessing rate from context,

meaning that the lexical guessing task results were better for the 96% lexical coverage version of the passage.

The study of Hu & Nation (2000) concluded that readers need at least a vocabulary of 5,000 word families to cover approximately 98% of the running words in a novel and to achieve adequate unassisted comprehension. Few L2 readers were able to comprehend well with around 90%-95% vocabulary knowledge in a given text.

In 1989, Koda examined the correlation between vocabulary tests and two reading tests. The result of her study showed that vocabulary knowledge contributes to students' reading comprehension in Japanese, and that there was a correlation coefficient of 0.74 between the vocabulary test scores and the comprehension test scores.

Determining the vocabulary size is important in ascertaining whether it enables the foreign language learner to reasonably comprehend the written texts. It is also crucial to know whether there is a vocabulary knowledge threshold below which reading comprehension cannot be achieved, a level the readers absolutely need in order to transfer their L1 reading strategies, especially when the L1 writing system is different than the L2 (as is the case in Arabic).

Vocabulary Knowledge in AFL

Arabic has a different writing system than the European languages. Arabic is written and read from right to left and consists of 28 basic letters (consonants) and three main vowels (long and short). When a short vowel comes at the end of

the word, it typically indicates the grammatical function of it. “Arabic verbs, nouns and adjectives are a combination of root letters and affixes, consisting of short and long vowels, and consonants” (Abu-Rabia, 2002, p. 300). The root is “this core group of consonants that gives the basic meaning to a family of words” (Al-Batal, et al., 2004, p. 133). The root is not a word, but just a group of consonants; the vast majority of roots are trilateral roots (composed of three consonants), but there are also two- and four- consonant roots, but they are relatively rare.

Most Arabic words consist of a root and vowels (and possibly affixes) to give specific meaning. Many words derive from a common root and tend to be related in meaning. For instance, by adding vowels, or affixes to the root (3 consonants letters) ‘k-t-b’ (to convey the idea of writing), we can derive many other words which are related in meaning. From ‘k-t-b’ we may derive words such as kutub (books), kaatib (writer), maktuub (letter), maktab (office), maktaba (library). As may be observed, all these words share the same consonants k, t and b and related to the meaning, “writing”.

This root system can make the meaning of some Arabic words clear, as it is easy for learners to know the meaning of a word in Arabic by looking at its basic root, rather than having to memorize words that do not have any obvious connection with the idea they represent.

Also, as mentioned previously, the short vowels indicate the grammatical function of the word in the sentence that might also help the readers recognize the meaning of some words, but most Arabic written texts, such as newspapers, are

not voweled, which sometimes makes discerning the word meaning a difficult task. The learners can learn to rely on the root system or the context to recognize the word meaning in the text. “The sentence context conveys the primary effect to facilitate word recognition when texts are presented unvowelized for skilled readers” (Abu-Rabia, 2002, p. 305).

Whether the Arabic text is voweled or not, AFL readers need to know a certain percentage of known words in texts that will help them recognize the unknown words in the texts and to adequately comprehend the written text.

Vocabulary Knowledge and Reading Comprehension in AFL

In the Arabic as a second language research, it would appear that the only study that has investigated the relationship between vocabulary knowledge and reading comprehension is Khalidieh (2001). Khalidieh found that vocabulary knowledge had a significant effect on reading comprehension of expository texts. He suggested that reading comprehension “depends mainly on vocabulary knowledge” (p. 416). His study showed a high correlation ($p < .0000$) between vocabulary scores and overall reading comprehension scores which suggests that overall reading comprehension is directly proportional to knowledge of vocabulary. In his opinion, reading comprehension depends largely on the knowledge of vocabulary rather than on grammatical rules. But still this study can be criticized for not defining the ratio of familiar to unfamiliar words needed for reading comprehension.

L2 Linguistic Threshold

The Definition of Linguistic Threshold

The basic definition of the word “threshold” is an entrance or doorway, or, more abstractly, a level, point, or value above which something is true or will take place and below which it is not or will not (Webster’s Ninth New Collegiate Dictionary, 2nd ed., s.v.” “threshold”) (Lee & Schallert, 1997).

Smith (1995) called a language threshold a “passport for learning” (p. 20). He declared that “you cannot become a member until you are familiar with the language, and you cannot learn the language until you join the club” (p. 20).

As mentioned in the previous chapter, Nation (2001) suggests two definitions of the lexical threshold. The first definition can be considered as the traditional meaning that sees a threshold as “an all or nothing phenomenon. If a learner has not crossed the threshold, then adequate comprehension is not possible”(p. 144). If a learner has crossed the threshold, comprehension becomes possible. The second definition of the language knowledge threshold, proposed by Nation, suggests that if a learner has not crossed the threshold, the possibility of comprehension is low: “if the learner has crossed the threshold, the chances are on the side of the learner gaining adequate comprehension.” This is what Nation described as a “probabilistic boundary” (p. 144).

The Development of the Notion of Linguistic Threshold

The notion of a second language threshold was first presented by Clarke (1980), when he suggested that insufficient second language competence results in “short circuit” transferring of L1 skills to L2. Cummins (1980) continued the research along similar lines. He presented the notion of a threshold for language proficiency, calling for thresholds of different skills of language knowledge. The research on the threshold hypothesis continued with a particular focus on a lexical threshold and its effect on L2 reading comprehension. The nature of the language threshold is largely lexical and the vocabulary knowledge threshold states that a reader must achieve a certain linguistic level in L2 before other variables, such as background knowledge and reading strategies, begin to affect reading comprehension (Ridgway, 1997).

Lexical Threshold for L2 Reading Comprehension

The idea of a threshold or linguistic ceiling in relation to reading comprehension suggests that L1 reading skills do not magically transfer from one language to another, especially when there are two writing systems involved (Koda, 1988).

Secondly, readers will not be able to read effectively until they develop some proficiency in the second language. Laufer (1992) agrees with the belief that there is a floor or threshold of competence below which it would be unreasonable to expect foreign readers to apply any reading strategies. In 1995, Bernhardt and

Kamil defined the lexical threshold: “in order to read in a second language, a level of second language linguistic ability must first be achieved” (p. 17).

The linguistic knowledge needed for reading comprehension is primarily lexical. “The nature of the threshold of reading comprehension indicates that vocabulary is the most critical element for interpretation of the threshold” (Huang, 1999). Laufer (1992) states that the most important issue is to find “the number of words the reader must possess in his lexicon to be able to read in L2, namely the number of words constituting the threshold vocabulary which will ensure the transfer of reading skills from L1 to the L2” (p. 127).

In 1985, Laufer and Sim used comprehension questions and interviews with L2 learners to determine the threshold scores of L2 reading competence. The results suggested that a minimum necessary comprehension score is 65% - 70% on the First Certificate English examination, below which L2 learners can't transfer their L1 strategies to L2. They suggest that vocabulary is the most important need for reading comprehension, then subject matter knowledge, and then syntactic structure.

In 1989, Laufer took a step further to see what percentage of word tokens (running words) needed to be known in order to ensure reasonable comprehension of the text. She set 55% to be a score for reasonable comprehension. The results show that the group who scored 95% and above on the vocabulary test had a significantly higher score (55% and above) in the reading test than those scoring below 95%.

Laufer (1992) correlated Israeli university students' vocabulary sizes with their reading comprehension scores, and found that the knowledge of 3,000 word families, which provide about 95% lexical coverage in a text, would result in reading scores of 56%. This is the lexical threshold for comprehending academic texts in English. The study showed that the ratio of known words and unknown words in the reading passages strongly affects the readers' comprehension.

Coady (1997) agreed that this 3,000-word family level is the minimum vocabulary knowledge that an L2 learner should have before reading strategies are effective. Also, in 1992, Hirsh & Nation compared the comprehension score to the proportion of words known in three short novels; the results of their study indicates that L2 learners need to have a vocabulary knowledge of around 5,000 word families, which is typically the 98% threshold of vocabulary coverage in a given text, in order to comprehend un-simplified texts and for pleasure reading.

Nation 2001 suggests that “any number of word families needed to cover certain percentage in a text depends on: 1) Type of text – novel, newspaper, academic text, etc., 2) Length of text, and 3) homogeneity of text; is it on the same topic and by the same writer?” (p. 146).

Also, it is worth mentioning here that reading comprehension of given texts also depends on other variables beside the lexical threshold, like the familiarity of the topic, or prior knowledge of subject matters. Carell & Eisterhold (1983) claim that some L2 readers' reading problems stem from insufficient background knowledge of the topic of the written texts. But L2 readers need to

achieve the linguistic threshold before background knowledge can affect L2 reading comprehension (Ridgway, 1997).

Based on the above literature, one can conclude that the linguistic threshold is mainly lexical for reading comprehension, that L2 learners need sufficient lexical coverage to achieve adequate comprehension. Some English as a Second Language (ESL) researchers agree that this 95% means that L2 learners should possess a working knowledge of around 3,000 word-families or more, based on the type of the test, as shown in previous studies.

Authenticity in L2 Reading

Definition of Authentic Materials

As mentioned in the previous chapter, the use of authentic material in L2 classrooms dates back to the 1970s, when the communicative language teaching (CLT) approach spread. Since then, the definitions of “authentic material” “tend to reflect the primacy of communicativeness” (Mishan, 2005, p. 12). Swaffar (1985) defines authentic material as the text that can be written to language learners, by native speakers, as long as there is an “authentic communicative objective in mind.” “One of which is written for native speakers of the language to be read by other native speakers or it may be a text intended for a language learner” (p. 17). It is real language created by native speakers of the target language in pursuit of communicative outcomes (Little, Devitt, & Singleton, 1989).

“A text is usually regarded as authentic if it is not written for teaching purposes but for a real-life communicative purpose, where the writer has a certain message to pass on to the reader. As such, an authentic text is one that possesses an intrinsically communicative quality” (Lee, 1995, p. 324).

In 1978, Widdowson defined authenticity by making a distinction between two terms, authentic and genuine texts. To him, the “genuineness is a characteristic of the passage itself, and is an absolute quality, while the authenticity is a characteristic of the relationship between the passage and the reader and it has to do with appropriate response” (Widdowson 1978, p. 80). (as cited in Mishan, 2005).

According to Widdowson’s definition, the genuineness of texts does not ensure authenticity, as authenticity has two poles, the characteristic of the text and the recipient of this authentic texts; how the reader will react to this authentic text, “without a high degree of relating the text to the reader/hearer, authenticity of the text cannot be said to have been realized” (Feng & Bryam, 2002, p. 59).

“Authenticity is not brought into the classroom with the materials or the lesson plan; rather, it is a goal that teacher and students have to work towards, consciously and constantly” (Van Lier, 1996, p. 128). Authenticity is not a quality of a text, rather than the way in which genuine texts are used by teachers and students. If they fail to authenticate a text by not responding to it as people do in non-instructional setting, not dealing with it in a way corresponds to normal communicative language, not using authentic tasks, or using inappropriate texts, then the authenticity of this text is not met (Mulling, 1991).

According to the above definitions, authenticity is seen to be related to the communicative intent and how the readers interact with these authentic materials to reach the communicative goal. Rather than defining authenticity, Mishan (2005) thinks that it is more important for language learning material developers to set criteria for authenticity that can be used for assessing the authenticity of the texts. She sets the following “criteria of authenticity”. “Authenticity is a factor of the:

- Provenance and authorship of the text.
- Original communicative and socio-cultural purpose of the text.
- Original context (e. g. its source, socio-cultural context) of the text.
- Learning activity engendered by the text.
- Learners’ perceptions of and attitudes to, the text and the activity pertaining to it” (p .18).

Comprehensibility of Authentic Materials in L2

When using authentic materials in the foreign language learning process, one issue that should be taken into consideration is the balance between using these authentic materials and how comprehensible they are. Authentic reading texts can be motivating to students, but also overwhelming. Some authentic texts may contain language that is beyond the learner’s abilities; in this case authentic materials will be frustrating rather than motivating (Bacon, 1992). In order to avoid that, FL teachers and materials developers need to determine what kind of authentic text is appropriate for what level and “what authentic texts are for a

particular interest to a class (poetry, newspaper articles)” (Swaffar, 1985, p. 18).

Instructors should choose texts that are at an appropriate level of linguistic difficulty, adapting authentic texts if needed (Day and Bamford, 1998).

Nuttall (1996) suggests three criteria for choosing reading texts for L2 learners, among them is the readability of the text. Texts should be comprehensible, but challenging to learners without overwhelming them (as cited in Brown, 2007). “Posing a reasonable challenge to the students - neither too difficult nor too easy” (Ellis, 1994, p. 516) - is important in language learning. It motivates L2 learners and “instilling a sense of achievement” (Mishan, 2005, p. 60). When learners realize they can successfully understand authentic texts, confidence in their own target language abilities increases (Leloup & Pontiero 1997).

In 1981, Krashen introduced the concept of “comprehensible input” and explained it in his famous formula (i+1) that suggests that the input (i) is comprehensible to the learner even if it is somewhat above his/her current proficiency level (+1), but he stresses that the “input” should be comprehensible enough for learners to comprehend meaning from the text and for language acquisition in general to take place.

Some researchers argue that this comprehensibility can be achieved by making some simplifications or modifications to authentic texts (Day & Bamford, 1998). Simplifying authentic texts might be the solution to ensure comprehensibility, but what the L2 learners might need in order to comprehend some authentic material, as suggested by Swaffar (1985), is to present a brief

target language introduction regarding some cultural components that might be obscure to them and this will avoid what he calls “spoon-feeding” simplified or edited texts. Bernhardt (1984) believed that these edited texts do not focus on the actual reading comprehension process and increase the difficulty of the texts, as the process might remove elements crucial to comprehension.

From the above discussion, one can conclude that the comprehensibility of authentic materials “still remains a central requisite for texts to promote language learning – a rich but comprehensible input of real spoken and written language in use” (Mishan 2005, p. 24).

In conclusion, the literature review indicates that there is a relationship between vocabulary knowledge and reading comprehension. Vocabulary knowledge has a great impact on reading comprehension. No text comprehension is possible, either in one’s native language or in a foreign language, without understanding the text’s vocabulary. When the percentage of unknown vocabulary increases, the possibility of comprehending the text decreases. L2 readers need sufficient vocabulary that acts as a lexical threshold between adequate and inadequate comprehension. L2 readers need to achieve this lexical threshold before other variables like background knowledge and L1 reading strategies can start to influence reading comprehension. Many ESL studies investigated the percentage of word coverage needed for reading comprehension in different type of texts. In 1992, Nation suggested 98% of known words for comprehension of Novels. Laufer (1985, 1989, 1992) found that 95% of known words are needed for reasonable comprehension of any authentic text.

In AFL research area, only one study investigated the relationship between vocabulary knowledge and reading comprehension, but none have investigated the percentage of word coverage needed to assure reading comprehension of Arabic written text.

These previous studies indicate that there is a need for this study to look closely at the relationship between vocabulary knowledge and reading comprehension of Arabic authentic texts. In particular, this research measures the correlation between the percentage of known vocabulary in authentic Arabic texts and the degree of comprehension of the text and whether there is a lexical threshold needed for AFL learners to comprehend authentic texts.

Based on the available research listed above, the following chapter addresses the methodology followed in this research to answer the research questions outlined in Chapter 1.

Chapter 3: Methodology

The present study attempts to measure the correlation, if any, between the percentage of known vocabulary in authentic Arabic texts and the degree of comprehension of the text. It also attempts to determine whether there is a lexical threshold for comprehending authentic Arabic material, below which learners cannot adequately comprehend the authentic texts. The rationale for undertaking this study is that no study has previously been done to measure the amount of vocabulary needed to read in Arabic as a second language. Previous studies in other languages, such as English suggest that second language readers need to understand 95% of word tokens in a text to ensure reasonable reading comprehension of the text (Laufer, 1989, 1992). Laufer found that when learners know about 3000 words in EFL, they will know about 95% of word tokens in any authentic text. She proposed that 3000 is the threshold level below which readers will not be able to adequately comprehend authentic material. Another study by Nation (2000) found that EFL learners need 98% of lexical coverage to comprehend un-simplified text for pleasure. This study is designed to investigate the following two research questions:

- Is there a correlation between the percentage of known vocabulary and reading comprehension of authentic Arabic texts?
- Is there a minimum percentage of known vocabulary that acts as a lexical threshold for AFL learners to effectively read Arabic news articles?

The Subjects

This section describes the criteria for choosing the subjects and provides demographic data.

Second-year AFL learners at Brigham Young University were chosen for this study. According to the ACTFL/ILR proficiency scale, they ranged from Intermediate Low to Intermediate Mid in both productive and receptive skills. The minimum amount of vocabulary knowledge was about 1600 words, based on the number of vocabulary they were exposed to and expected to learn during first- and second-year Arabic classes from their text books *Al-Kitaab fii Ta'allum al-Arabiyya – Part One & Two*. In addition, they were exposed to many authentic short news articles and the Egyptian dialect, providing them with additional vocabulary. This vocabulary is not included in 1600.

The subjects' participation in this study was on a voluntary basis. They were informed two weeks in advance of the tests about the study. On the cover of each test there was a consent form meant to encourage the students to participate in this study. Only those that read and agreed to it proceeded to take the test.

The total number of subjects is 23, with 18 males (78%) and 5 females (22%). The average age is 23.3 years, with an age range of 18 to 28. As shown in

Table 1, the majority of the subjects (74%) are between the ages of 23 and 28. All participants are native English speakers, with almost all of them, i.e. 22 (96%), possessing a knowledge of at least one foreign language, and 13 (57%) subjects having knowledge of two besides Arabic (Table 2).

Table 1 : Age Range of Subjects.

Age Range	Number of Subjects	Percentage
18-22	6	26
23- 28	17	74

Table 2: L2 Background Knowledge (Excluding Arabic).

L2	Number of Subjects	Percentage
One L2	22	96%
Two L2	13	57%

Material

The reading materials consisted of two authentic passages averaging 143 words taken from an Arabic online news source. These texts were not written for people with a limited vocabulary as they contained many words unfamiliar to learners. These two passages were not randomly selected. They were chosen based on the ILR standards for rating reading passages and with the goal of trying to match them with AFL learner's proficiency levels. Passage one is at the same level of the subject's reading proficiency level (Intermediate) and is a biography of the Egyptian Nobel Prize recipient, Dr. Ahmad Zewail. Subjects had already been exposed to short biographies in their Arabic textbook, so they had

experience in reading this type of article. Passage two is harder than the first passage. According to the ILR scale, it is one level above the subjects' proficiency level. It is a political news article about the Palestinian-Israeli conflict. This kind of political news articles is not new to the subjects, as they practiced with many such short news articles.

Instruments

Background Questionnaire

This questionnaire given to the subjects was designed to gather background information on the subjects' genders, ages, L1 and L2 background knowledge (duration and place of learning), and time spent learning Arabic.

Reading Comprehension Test (RCT)

In this test, open-ended questions (constructed- response questions) were used. This type of question has been used by many researchers in assessing learners' reading comprehension (Alderson et al., 1995; Nation & Hu, 2000; O'Malley & pierce, 1996). This kind of question gives the students a better opportunity to show their understanding of the details of the text (Nation & Hu 2000) and to construct their own answers (Abanomey, 2002), especially when compared to multiple-choice questions. Open-ended questions "call for the examinee to produce something instead of merely choosing between two or more alternatives" (Popham 1990, p. 248).

However, one of the disadvantages of open-ended questions, as seen by some researchers, is the difficulty of scoring them satisfactorily, as they are affected by subjective judgments of the evaluators (Badger & Thomas, 1992).

The test consisted of five open-ended questions that covered almost all of the details given in the text. All questions were written in English and subjects were asked to answer in English as well, since the main purpose was not to test their understanding of the questions but their comprehension of the text. Each question in the test was scored on a four-point scale. The highest possible score was 20 points. Another evaluator was asked to grade the test to ensure the validity of scoring and the independence of results.

To determine the minimum score required for adequate comprehension of the texts, Nation's (2000) method was used to decide what should be considered as adequate comprehension of the text, which is to use an arbitrary decision that allows a degree of human error. To apply this method on this study, 20 out of 20 is obviously adequate comprehension and if we allow for human error, then 15 out of 20 (75%) should also be acceptable. Accordingly, the decision was made to consider 15 out of 20 the minimum score of adequate comprehension.

Identifying Unknown Vocabulary

The participants were asked to circle all unknown or unfamiliar vocabulary in the passage after attempting to comprehend as much of the passage as possible; this task was used previously by Laufer (1989) to serve the same

purpose as the present study. This was required to measure the percentage of unknown vocabulary.

Lexical Coverage Test (LCT)

After circling the unknown vocabulary, the subjects were given a Lexical Coverage Test on the passage based on their circling of unknown vocabulary. This LCT included an average of 40 words, which the students were asked to translate or paraphrase the meaning. The reason behind this task was to get an accurate list of unknown vocabulary, by adjusting the number of words mistakenly left uncircled but still unknown as determined by a lexical coverage test (Laufer 1989, Nation 2000). Nation considers this test much more demanding and accurate than a Multi-Choice Recognition Test, as it eliminates the option of randomly guessing meanings from the choices presented.

Based on the Lexical Coverage Test, there are four possible options regarding word familiarity for each word:

1. A word was not circled and was correctly translated in LCT (considered a familiar word).
2. A word was circled but left un-translated or mistranslated (unfamiliar).
3. A word was mistranslated or left un-translated, and yet was not circled as unknown (unfamiliar).
4. A word was correctly translated but randomly circled as unknown (familiar).

Options 2 and 3 are considered to be the unknown words in the text, and options 1 and 4 are the known words used to calculate the percentage to be correlated with the reading comprehension test scores.

Data Collection Procedures

The data was collected through two tests for each reading passage: RCT and LCT, and a circling task. Each set of tests was collected in two separate sessions. The students, who voluntarily participated in this study, were asked to take a total of four sessions for the two reading passages. One week separated these sessions. It took about four weeks to collect all the required data. The scheduled two sessions were administered as follows:

Session One

In this session, the participants took the reading comprehension test and proceeded with the circling task. The students were not limited by time in the reading comprehension tests, which involved reading the passage as many times as needed to confirm whether they comprehended it before answering the open-ended questions in English. The majority of the subjects spent about one hour and a half to complete this test. In the circling task, the students circled all the words unknown or unfamiliar to them. This task took about fifteen minutes, so the total amount of time spent in session one was one hour and 45 minutes.

Session Two

To prepare for the second stage of the study, the circling task data was used to create a list of words to help design a lexical coverage test (LCT) based on the words circled by the students. The subjects were asked to re-read the passage and translate the list of vocabulary in their L1. The subjects took about 45 minutes to finish this test.

Research Design

All the research data was collected through the background questionnaire (Appendix A), reading comprehension test (RCT), circling of unknown vocabulary task (Appendix B, C), and lexical coverage test (Appendix D, E). In order to investigate the effect of the percentage of vocabulary coverage on reading comprehension in Arabic for foreign language learners, the research design was based on the following null hypotheses.

Null Hypotheses

Based on the two research questions, this study sets up the following null hypotheses:

- H₀ 1: There is no correlation between the percentage of known vocabulary and reading comprehension of authentic Arabic texts.
- H₀ 2: There is no minimum percentage of known vocabulary that acts as a lexical threshold for AFL learners to effectively read Arabic news article.

Variables

The major purpose of this study was to investigate the effect of the percentage of vocabulary coverage (known words) on the reading comprehension of authentic Arabic texts. This study involved one dependent variable, the reading comprehension score (scores in two Arabic reading comprehension tests); and one independent variable (predictor variable), coverage of known vocabulary. In order to examine the correlation between the dependent and independent variable in this study, the linear regression was used to analyze the collected data.

Dependent Variable

The dependent variable of this study was the reading comprehension scores. This variable was measured using the scores collected from the two Arabic reading comprehension tests. The scores of these two tests were taken from answering five open-ended questions for each passage after the subjects read the Arabic reading passage. These five open-ended questions covered all details of the passage and each question was worth four points with a total possible score of 20 points.

Independent Variable

The independent variable of this study was the percentage of vocabulary coverage (known words). Each subject was asked to circle all unknown words in each text. The total number of unknown words for each subject-passage combination was adjusted based on the Lexical Coverage Test. The adjustment is

necessary because some of the words were mistakenly considered as known. Thus, the total number of unknown words is the total number of circled words plus the number of unknown words determined by the Lexical Coverage Test. This constitutes the total number of unknown words. The total number of known words is then determined by subtracting the total number of unknown words from the total number of words in the passage. The percentage of known words is thus determined.

Statistical Analysis

In order to analyze the dependent and independent variable separately, descriptive statistics were used to determine the means and standard deviations of the reading comprehension scores (dependent variable) and percentage of known words (independent variable).

This study tested what percentage of word knowledge AFL intermediate learners need in Arabic authentic texts to achieve adequate comprehension. Simple linear regression was used to correlate the dependent and independent variables of this study.

It was necessary to test the validity of the statistical analysis used in this study and to see if the test assumptions are not violated. The normality assumption is a key assumption for the simple linear regression. Therefore, a normal probability plot was derived for both passages.

To summarize, this study explores the relationship between the vocabulary knowledge and reading comprehension of Arabic authentic texts. It measures the

correlation, if any, between the percentage of known vocabulary in authentic Arabic texts and the degree of comprehension of the text. It also measures the lexical threshold necessary for comprehending authentic Arabic materials, below which learners cannot adequately comprehend authentic texts.

Twenty-three intermediate AFL learners at Brigham Young University participated in this study. They were given two reading comprehension tests, circling the unknown word task, and lexical coverage test for each passage. The two authentic texts were chosen from Arabic online news source. The reading comprehension scores is the dependent variable and percentage of known words is independent variable of this study. In order to know the achievement of each variable, Descriptive statistics was used to determine the means and standard deviations of the reading comprehension scores and the percentage of known words. Simple linear regression was employed to correlate the dependent and independent variables of this study.

The following chapter utilizes the methodology outlined in this chapter and lists the main research results that answered the research questions.

Chapter 4: Results and Discussion

The purpose of this study is to investigate the relationship between vocabulary knowledge and reading comprehension of authentic Arabic texts. Specifically, it tests the percentage of known words (vocabulary coverage) that the AFL intermediate learner needs to adequately comprehend two Arabic authentic texts (news articles). The study also examines whether there is a minimum threshold of vocabulary knowledge for reading comprehension of authentic text in Arabic as foreign language.

As mentioned in Chapter 3, this study includes one dependent variable, the reading comprehension score (scores for two Arabic reading comprehension tests) and one independent variable (predictor variable) which is the percentage of known vocabulary. In order to examine the correlation between the dependent and independent variables in this study, the simple linear regression was used to analyze the collected data.

The study tried to answer the following research questions:

- Is there a correlation between the percentage of known vocabulary and reading comprehension of authentic Arabic texts?
- Is there a minimum percentage of known vocabulary that acts as a lexical threshold for AFL learners to effectively read Arabic news articles?

Based on the above research questions, the following null hypotheses were formulated:

- H₀ 1: There is no correlation between the percentage of known vocabulary and reading comprehension of authentic Arabic texts.
- H₀ 2: There is no minimum percentage of known vocabulary that acts as a lexical threshold for AFL learners to effectively read Arabic news articles

Mean and Standard Deviation of the Dependent Variable

As mentioned earlier, the dependent variable of this study is the reading comprehension score, which includes the comprehension scores from reading the two Arabic passages and answering five open-ended questions about each passage. The total possible score of the test was 20 points, four points for each question. Table 3 shows the mean and standard deviation of the two passages' reading comprehension scores. The mean score for reading passage one is 16 with a standard deviation of 3.86. For reading passage two, the mean score is considerably lower, i.e. 9, with a standard deviation of 6.23.

Table 3: Mean & SD of Reading Comprehension Scores.

Reading Comprehension Scores	Mean	Standard deviation
Passage 1 (biography article)	16	3.86
Passage 2 (political article)	9	6.23

Figure 1 and Figure 2 show the distribution of the reading comprehension test scores in reading passages one and two respectively.

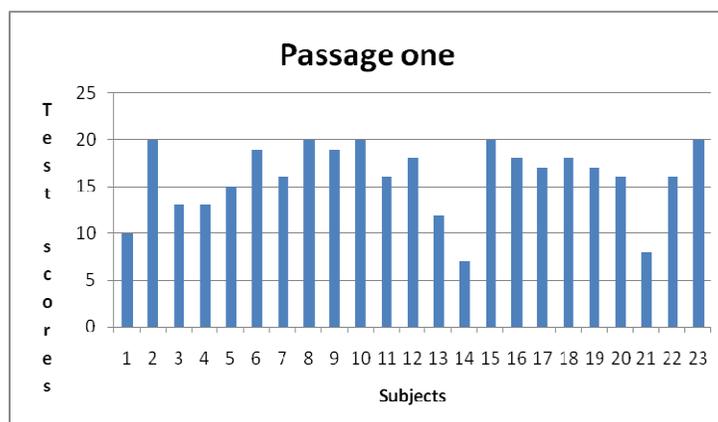


Figure 1: RCT Scores – Passage 1.

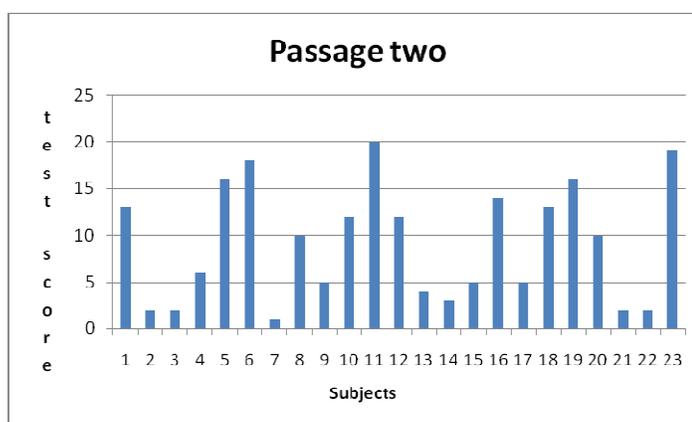


Figure 2: RCT Scores – Passage 2.

As seen in Figure 1, five subjects (22%) scored 20 points out of 20; twelve subjects (52%) received a score of 15 or higher. Most subjects (74%) had an adequate comprehension of the text. In Figure 2, only one subject scored 20 out of 20, and only four subjects (17%) scored 15 or higher. Also, seven subjects scored less than five points. This variation in test scores in passage two resulted in lower mean score and higher deviation than passage, as seen in Table 3.

Mean and Standard Deviation of the Independent Variable

The independent variable of this study is the percentage of known vocabulary in the reading passages. As mentioned in the previous chapter, this variable was measured by the students' circling of unknown vocabulary and by the lexical coverage tests. Table 4 shows the mean and standard deviations of the percentage of known vocabulary in both reading passage one and two.

Table 4: Mean & SD of Word Coverage Percentage.

Percentage of Word Coverage	Mean	Standard Deviation
Passage 1	91.3%	4.34
Passage 2	77%	8.40

As shown in Table 4, the average percentage of known words among subjects in passage one is 91.3% with a standard deviation of 4.34. For passage two, the average percentage of known words was 77%, with a standard deviation of 8.4. As noted, the mean percentage of known words in passage one is higher than passage two, and this suggests that passage two was more difficult for the subjects than passage one and contains more words that are unfamiliar to them.

Figure 3 and Figure 4 show the distribution of percentage of known words in passage one and two respectively. By comparing Figure 3 and Figure 4 we can notice that most of the subjects know more words in passage one, than in passage two and this affected their scores in the comprehension tests.

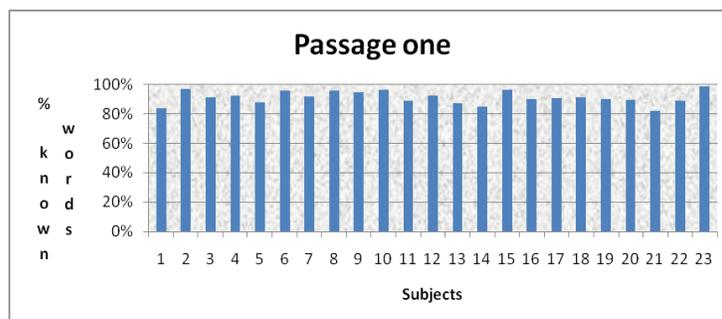


Figure 3: Distribution of Known Words Percentage among Subjects - Passage 1

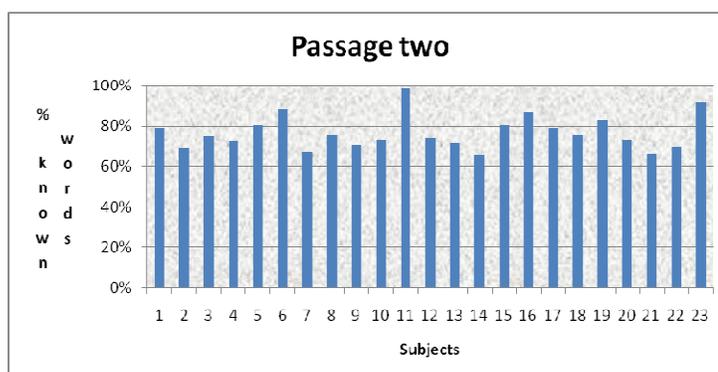


Figure 4: Distribution of Known Words Percentage among Subjects - Passage 2

Validity of the Statistical Procedures

As indicated in Chapter 3, it was necessary to test the validity of the statistical analysis and see if the statistical test assumptions were violated. The normality assumption is a key assumption for the simple linear regression. Therefore, a normal probability plot is derived for the regression analysis for both passages (see Figure 5 and Figure 6).

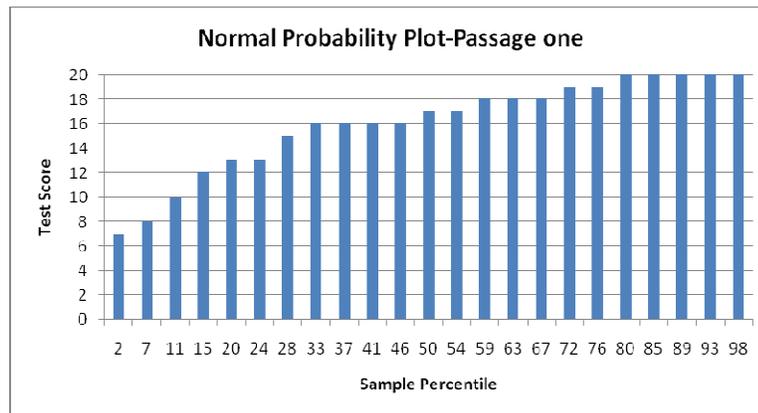


Figure 5: Normal Probability Plot – Passage 1.

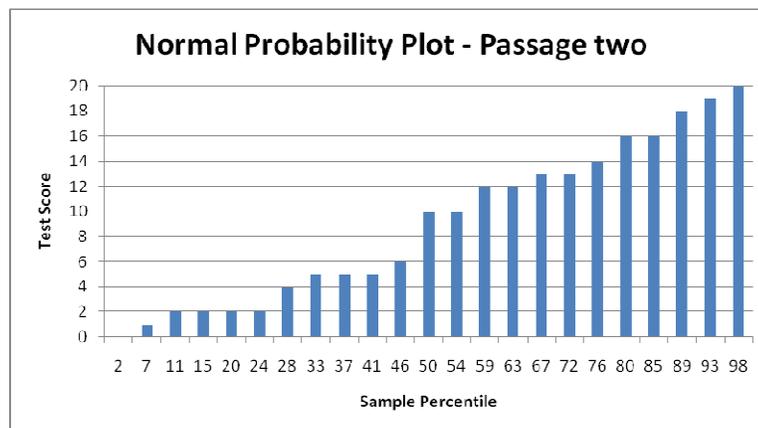


Figure 6: Normal Probability Plot – Passage 2.

As we can see from Figure 5 and Figure 6, both plots are almost linear, which may indicate that the normality assumption is not severely violated. On the other hand, the t-tools, used within the regression analysis, are robust against slight deviations from normality. Hence, the available data indicates that the t-tools and regression analysis may be used for the construction of a predictive

model. The confidence interval summarizes uncertainty that comes from errors in the whole process of the survey.

Variables Correlation

A simple linear regression was used to determine if there is a correlation between the reading comprehension scores in the two Arabic reading passages and the percentage of known vocabulary. Figure 7 and Figure 8 show a graphic representation of the least squares regression line. In these figures, the correlation coefficient, between the percentage of known vocabulary and the reading comprehension scores in passage one and two, is shown.

Passage One

The statistical analysis of passage one (the biography passage) indicates that there is strong evidence to show a relationship between the percentage of known words and test scores. The p-values as seen in Table 5 for both the slope and intercept are significantly low ($p < .0000$). Also, the t-statistic for the intercept and slope are significantly high. This indicates that both the intercept and the slope are key factors in predicting the subjects' test scores (the dependent variable) based on the percentage of known words (the independent variable). Another evidence of the relation is that none of the slope and intercept 95% confidence intervals contains zero. It must be mentioned here, however, that these inferences are only valid for the study at hand and are specific to this study's subjects (AFL intermediate learners).

Table 5: Regression Analysis Results – Passage 1.

	Coeff.	St. Error	t-Stat	P-value	Lower 95%	Upper 95%
Intercept	-54.66	8.77	-6.22	0.00000353	-72.92	-36.41
% Known Words	77.39	9.60	8.05	0.00000007	57.42	97.37

A correlation coefficient explains how much association is represented by the regression line and the dataset in hand. A relatively high correlation coefficient of 0.75 (Figure 7) may be seen as an indication of the regression line being a good representation of the dataset. Hence, it could be used for prediction. A positive value for the correlation coefficient indicates that comprehension test scores increase with an increase in percentage of known words.

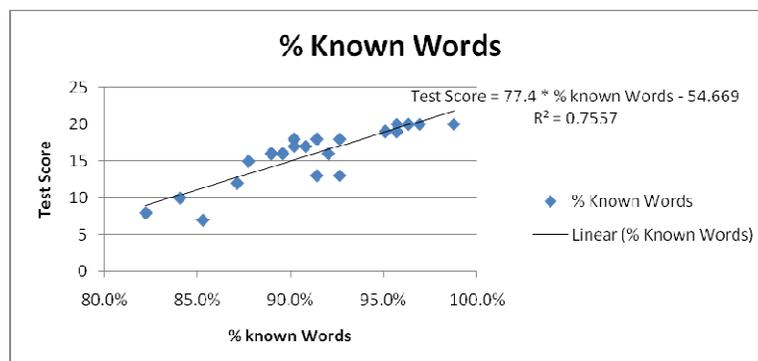


Figure 7: Percentage Known Words vs. Comprehension Score – Passage 1.

The regression line has a slope coefficient of 0.77 (with a 95% confidence interval ranging between 0.57 and 0.97). This may be viewed as follows: for every 1% increase in known words, there is a 0.77 increase in the comprehension test score (with a 95% confidence interval ranging between 0.57 and 0.97 points).

The regression line has an intercept of -54.67 (with a 95% confidence interval ranging between -36.4 and -72.9). This may be viewed as a threshold to adjust test scores after accounting for the regression line slope.

As mentioned above, the minimum test score for adequate comprehension of the texts was set to be 15 out of 20 points (75%). Table 6 below shows the minimum percentages of known words needed for AFL intermediate learners to comprehend this specific passage, or, in other words, to gain a score of adequate comprehension. The statistical analysis shows that the subjects need to know at least 90% of the text's words to adequately comprehend the passage. As shown in Table 6, the highest comprehension score at 20 points (100%) and the lowest score at 0 points (0%) and the minimum percentage of known words associated with all were calculated. The data shows that, in order to reach the highest comprehension level, which is 100%, the subjects need at least 96.5% of known word in this reading passage. Also the table suggests that a minimum of 70% of word coverage does not guarantee any comprehension. The scores in between the maximum and minimum scores are also measured as seen in Table 6.

Table 6: Minimum Percentage of Known Words – Passage 1.

Test Score		Min % of known words		
Out of 20	%	95% C.I. Lower Limit	Mean	95% C. I. Upper Limit
20	100	100%	96.5%	57.9%
15	75	100%	90.0%	52.8%
10	50	100%	83.6%	47.7%
5	25	100%	77.1%	42.5%
0	0	100%	70.6%	37.4%

Passage Two

As for Passage two (about the Israeli-palestinian conflict), the statistical analysis for the second passage indicated that there is a correlation between percentage of known words and test scores. The p-values for both the slope and intercept are significantly low ($p < .0000$) (see Table 7). Also, the t-statistics for the intercept and slope are high. This indicates that both the intercept and the slope are key factors in predicting test scores (the dependent variable) based on percentage of known words (the independent variable). Also none of the slope and intercept 95% confidence intervals contains zero, another key factor in the predictive equation. As mentioned previously, these inferences are only valid for the study at hand and are specific to AFL intermediate learners.

Table 7: Regression Analysis Results – Passage 2.

	Coeff.	St. Error	t-Stat	P-value	Lower 95%	Upper 95%
Intercept	-39.45	6.78	-5.81	0.0000090	-53.56	-25.33
% Known Words	63.13	8.79	7.18	0.0000004	44.85	81.41

A correlation coefficient explains how much association is represented by the regression line and the dataset in hand. A relatively high correlation coefficient of 0.69 (Figure 8) may be seen as an indication of the regression line being a good representation of the dataset. Hence, it could be used for prediction. A positive value for the correlation coefficient indicates that comprehension test scores increase with an increase in the percentage of known words.

As shown in Table 7, the regression line has a slope coefficient of 0.63 (with a 95% confidence interval ranging between 0.44 and 0.81). The regression line has an intercept of -39.45 (with a 95% confidence interval ranging between -25.3 and -53.5). This may be viewed as a threshold to adjust test scores after accounting for the regression line slope.

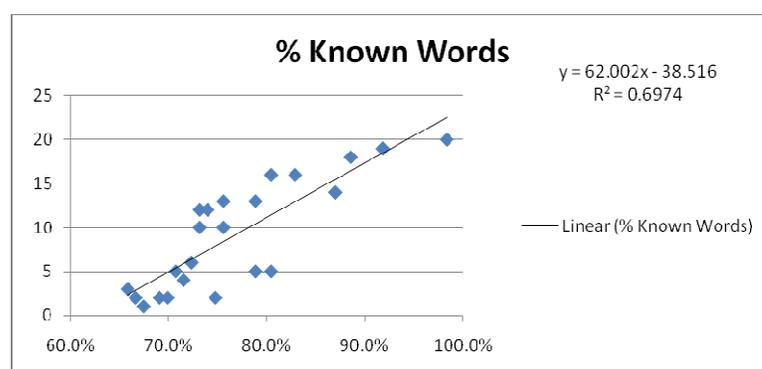


Figure 8: Percentage Known Words vs. Comprehension Score – Passage 2.

As mentioned above, the minimum test score for adequate comprehension of the texts was set to be 15 out of 20 points (75%). Table 8 below shows the minimum percentage of known words in this passage needed for AFL intermediate learners to comprehend this specific passage, in other words, to gain the comprehension test score that was previously set to be the minimum score for adequate comprehension. The statistical analysis shows that the subjects need at least 86.2% of word coverage to adequately comprehend the passage. As shown in Table 8, we measured also the highest comprehension score, which is 20 points (100%) and the lowest score 0 point (0%) and the minimum percentage of known

words associated with it. The data showed that in order to reach the highest comprehension level (100%) the subjects need at least 94.2% of word coverage in this reading passage. Also the table suggests that the minimum percentage of 62.5% of word coverage does not guarantee any comprehension for this reading passage. The scores in between the maximum and minimum scores are given in Table 8.

The main reason that the correlation coefficient value in Figure 8 is a little less than in Figure 7 is that 10 subjects (almost half) scored 5 or less out of 20. This can be attributed to the actual variations between L2 learners and the fact that individual performance may vary from a test/passage to another. This performance may be explained by other explanatory variables as mentioned earlier.

Table 8: Minimum Percentage of Known Words – Passage 2.

Test Score		Min % of Known Words		
Out of 20	%	95% C.I. Lower Limit	Mean	95% C. I. Upper Limit
20	100	100%	94.2%	55.7%
15	75	100%	86.2%	49.5%
10	50	100%	78.3%	43.4%
5	25	100%	70.4%	37.3%
0	0	100%	62.5%	31.1%

After statistically analyzing the data for each passage, the average of the percentages of known words needed adequate comprehension of both passages was given in Table 9.

Table 9: Minimum Percentage of Known Words – Mean, Both Passages.

Test Score		Min % of Known Words (Mean)		
Out of 20	%	Passage 1	Passage 2	Average
20	100	96.5%	94.2%	95.4%
15	75	90%	86.2%	88.1%
10	50	83.6%	78.3%	80.9%
5	25	77.1%	70.4%	73.8%
0	0	70.6%	62.5%	66.6%

As shown in Table 9, in passage two the subjects needed 86.2% as a minimum percentage of known words to score 15 (the minimum test score set for adequate comprehension in this study) or higher. This less than the minimum percentage of known words needed for passage one.

This may appear contradicting to the previous statement that passage two is harder than passage one, however, we have to understand that passage two has higher standard deviation than passage one and the mean of test scores and word coverage is almost half of passage one, and this indicates that there is higher variations among subjects in passage two and this can be explained by the larger deviations in both test score and percentage of known words (see Table 3 and Table 4). Its 95% confidence interval is wider than passage one. (see Table 5 and Table 7). Also, as seen in Figure 8, the correlation in passage two is less than in passage one, which indicates that this variation among the students affects the predictive equation.

Density and Coverage

In 2000, Nation & Hsueh looked at the issue of determining the amount of vocabulary needed by L2 learners to read with reasonable comprehension, from a “commonsense view” (p. 405) by arguing that the density of unknown word tokens and vocabulary coverage are related in various types of texts (such as novels, news articles). Based on their approach, the density of unknown words in relation to the percentage and amount of word coverage in the two passages was calculated (Table 10).

Table 10: Density of Unknown Words to Total Words.

Density of Unknown Words tokens	Known Word Coverage		
	Percentage	Passage 1 (163 Words)	Passage 2 (123 Words)
1 unknown to 99.0 known	99%	161	122
1 unknown to 50.0 known	98%	160	121
1 unknown to 32.0 known	97%	158	119
1 unknown to 23.4 known	96%	156	118
1 unknown to 19.0 known	95%	155	117
1 unknown to 9.00 known	90%	147	111
1 unknown to 7.30 known	88%	143	108
1 unknown to 4.00 known	80%	130	98

Table 10 shows that for passage 1, a subject needs to know 161 words out of the passage total word count, i.e. 163, to have 99% known word coverage for

this passage (161/163). In other words, this subject would know 99 words for every 100 words he reads, missing only one word.

For passage 2, a subject needs to know 122 words out of the passage total word count to have 99% known word coverage for this passage (122/123).

For a percentage of known words of 80%, a subject needs to know 130 words out of the 163 in passage 1 and 98 words out of 123 in passage 2. This means he would know 4 words and miss 1 of every 5 words in the passage.

Table 10 shows that the change in word coverage makes a change in the density of the unknown words. For example, if the subject does not know one word and know 99 words, this means a total number of words of 100 leading to 99% coverage (the first row in Table 10). Similarly, if a subject does not know one word and knows 19 words, this means that he knows 19 words in 20 (95%). With 98% coverage, the density drops to not knowing one in fifty words; and 97% coverage gives one in thirty-two.

It should be mentioned that this density and coverage differ from one kind of text to another. Since the above density and coverage was based on two reading passages from the media, we cannot automatically apply the same percentage to different texts and text types, but this study results can be used as an indicator that there may be a lexical threshold for reading comprehension below which the chances of adequate comprehension are low and that there is a certain percentage of vocabulary knowledge needed for AFL learners to comprehend Arabic authentic texts, as in the case of the two news articles in this study.

As mentioned earlier, this study sets the following two null hypotheses:

- H₀ 1: There is no correlation between the percentage of known vocabulary and reading comprehension of authentic Arabic texts.
- H₀ 2: There is no minimum percentage of known vocabulary that acts as a lexical threshold for AFL learners to effectively read Arabic news article.

The results of this Study provides enough evidence to reject the two null hypotheses set forth in this research. In passage one, there is a correlation coefficient of 0.7 between the percentage of known vocabulary and reading comprehension, and in passage two, the correlation is 0.6. The study suggests that there is a minimum percentage of known vocabulary that acts as a lexical threshold for AFL learners to effectively read Arabic news articles. In passage one, the learners needed 90% of known words to read the minimum score (15 out of 20 points) for adequate comprehension and those who have only 70% of known words showed no comprehension scoring zero out of 20 points. In passage two, the subjects needed around 86% to adequately comprehend the passage and a percentage of 62% failed to guarantee any comprehension of the passage.

Chapter 5: Conclusions

Most, if not all, researchers agree that vocabulary knowledge has a great effect on L2 reading comprehension. Many others feel that there is a threshold of vocabulary knowledge below which L2 reader cannot achieve an adequate comprehension of written texts. Many studies have been conducted to investigate this issue. Laufer (1989) found that readers need to know at least 95% of the words in a text for adequate comprehension of English academic texts. Hirsh and Nation (1992) determined that an un-simplified text can be comprehended when 95% of words are known. This acts as a lexical threshold below which L2 reader might not adequately comprehend the text. In 2000, Hu & Nation concluded that 98% of the words given in a reading text for pleasure are needed for adequate unassisted comprehension.

Few studies have investigated the relationship between vocabulary and reading comprehension in Arabic as a foreign language. These studies have not investigated the percentage of word coverage needed to assure reading comprehension. According to the available literature reviewed in this research, a lexical threshold below which the L2 learner cannot achieve the minimum level of reading comprehension was not yet established.

This research explored the relationship between vocabulary knowledge and reading comprehension of authentic Arabic texts. Particularly, it attempts to identify the percentage of word coverage readers need to ensure reading comprehension of two reading passages from online Arabic news source. It hypothesized that there is a threshold of vocabulary knowledge for reading comprehension for intermediate AFL learners to effectively read Arabic news articles.

The results indicate that there is a threshold of vocabulary knowledge for reading comprehension for intermediate AFL learners to effectively read Arabic news articles. Comparing the means of the comprehension scores and the vocabulary knowledge showed that the subjects who have word coverage of 90% or more in the first passage scored 15 out 20 points in the comprehension test. Below this percentage, the subjects could not reach adequate comprehension.

In passage two, the minimum percentage of word coverage needed for adequate comprehension was 86%. This indicates that there is strong evidence that the two null hypotheses set forth in this research are to be rejected. In that sense, it is evident that there is a direct relationship between the percentage of known words and the level of comprehension. These results also provide evidence that there is a minimum threshold of known vocabulary for the adequate comprehension.

Limitation of the Study

This study indicates that there is certain percentage of known words needed for AFL learners to adequately comprehend Arabic authentic texts and that there is a lexical threshold, below which AFL learners will not be able to achieve any reasonable comprehension and that learners need to reach this threshold before other variables like background information or L1 reading strategies can be effective. However, the results cannot be generalized and further studies should be conducted to validate the results for several reasons:

1. The small sample size. This study only used two reading passages from online Arabic news source, a very small corpus.
2. The subjects who participated in this study were only 23, which is not enough to confidently generalize the results.
3. This study did not account for differences that subjects might have regarding the familiarity of the topic that might affect their test scores and their vocabulary knowledge.

Also the validation of the comprehension tests was one of the limitations of this study. Some of the open-ended questions in passage two were formulated in a way that if a subject had a good background in the topic, s/he might be able to answer the questions correctly without actually comprehending the passage, which constitutes a threat to external validity. Finally, the test was administrated at the end of winter semester, so some subjects were stressed and might not have performed as well on the test as they could have.

Implications of the Study

The present study investigated the relationship between vocabulary knowledge and reading comprehension of Arabic authentic texts for AFL intermediate learners. Results show that there is correlation between percentage of known words and reading comprehension and that there may be a lexical threshold below which subjects cannot achieve adequate comprehension. Based on these results, some instructional implications are suggested for AFL teachers.

Implications for AFL Instructors and Teachers:

The results showed the importance of vocabulary knowledge in AFL reading comprehension of authentic texts, so AFL teachers should focus more on teaching vocabulary that is frequent in Arabic authentic texts.

Also the study suggested that there was minimum percentage of known words needed for AFL to comprehend the two reading passages, and that there was a lexical threshold below which subjects could not adequately comprehend the passages, so AFL teachers need to choose appropriate authentic texts according to the proficiency level of AFL learners, texts that do not have too heavy of a vocabulary load, challenging, but not overwhelming.

The findings of this study also suggests that AFL learners cannot completely cover all the words in Arabic texts, so teachers should teach reading strategies to help learners cope with unknown vocabulary and to apply their prior knowledge to guess the meaning from context.

Recommendations for Future Research

No research can investigate all variables that affect L2 reading comprehension. This study is just a step for further research, presenting the following recommendations for more research to validate and expand the results of this study.

As mentioned above, other variables may contribute to reading comprehension of AFL texts, so further research should add more independent variables, beside vocabulary knowledge, such as background information and reading strategies to investigate their affect on reading comprehension. This study investigated only the correlation between the percentage of known words and the degree of comprehension of intermediate AFL learners to effectively read Arabic news articles, so further studies should be conducted with AFL learners at different levels of proficiency, using different types of authentic texts. Also the number of subjects in this study was not large, so future study with larger numbers of subjects might give us clearer results. Also two passages are not enough of a sample to confidently generalize the results; larger samples might help to set a lexical threshold for AFL. The passages chosen for this study were not randomly selected, so it is recommended to replicate this study with randomly selected news articles and use standardized and valid comprehension tests. It is also important that the test/questionnaire designed in such a way that the scores will best represent subjects' comprehension.

In AFL is a lack of frequency lists that could guide teachers to use in choosing appropriate reading materials, to check if certain words are high

frequency words that L2 learners need to acquire to adequately comprehend authentic texts. Also vocabulary tests are very valuable instruments for AFL teachers and researchers to measure the learner's vocabulary size, so AFL vocabulary levels test should be designed. Useful research could further investigate the amount of vocabulary AFL Intermediate learner should possess to reliably read a text from any general domain.

Bibliography

- Abanomey, A. (2002). *The effect of texts' authenticity on reading-comprehension test-taking strategies used by adult Saudi learners of English as a foreign language*. Unpublished doctoral dissertation, Arizona State University.
- Abu-Rabia, S. (2002). Reading in a root-based morphology language: the case of Arabic. *Journal of Research in Reading*, 25(3), 299-309.
- Akagawa, Y. (1995). *The effects of background knowledge and careful attention on reading comprehension and vocabulary acquisition*. Unpublished doctoral dissertation, Temple University, Philadelphia, PA.
- Al-batal, M., Brustad, K., & Al-Tunisi, A. (2004). *Al-kitaab fii Ta 'allum al- 'Arabiyya, A Textbook for Beginning Arabic*. Washington, D.C.: Georgetown University Press.
- Alderson, J. C., Clapham, C, & Wall, D. (1995). *Language Test Construction and Evaluation*. Cambridge: Cambridge University Press.
- Alderson, J. C. (2000). *Assessing Reading*. Cambridge: Cambridge University Press.
- Anderson, R. C., & Freebody, P. (1981). Vocabulary knowledge. In J.T. Guthrie (Ed.), *Comprehension and Teaching: Research Reviews*. Newark, DE: International Reading Association.
- Bacon, S. M. (1992). Phases of listening to authentic input in Spanish: A descriptive study. *Foreign Language Annals*, 25, 317-334.
- Belnap, R. K. (2006). A profile of students of Arabic in U.S. universities. In K. M. Wahba, Z. A. Taha, and L. England (Eds.), *Handbook for Arabic Language Teaching Professionals*, (pp. 169-78). Mahwah, New Jersey: Lawrence Erlbaum.

- Badger, E., & Thomas, B. (1992). Open-ended questions in reading. ERIC/TM Digest.ED355253. *ERIC Clearinghouse on Tests, Measurement, and Evaluation*, Washington, DC.
- Barnett, M. A. (1989). *More than Meets the Eye: Foreign Language Reading: Theory and Practice*. Englewood Cliffs, New Jersey: Prentice Hall Regents.
- Bernhardt, E. (1984). Toward an information processing perspective in foreign language reading. *The Modern Language Journal*, 86 (4), 322-331
- Bernhardt, E. (1991). *Reading Development in a Second Language: Theoretical, Empirical, and Classroom Perspectives*. New Jersey: Ablex.
- Bernhardt, E. B., & Kamil, M. L. (1995). Interpreting relationships between L1 and L2 reading: consolidating the linguistic threshold level and the interdependence hypotheses. *Applied Linguistics*, 16, 15-34.
- Bossers, B. (1991). On thresholds, ceilings and short-circuits: the relation between L1 reading, L2 reading and L2 knowledge. In J. H. Hulstijin & J. F. Matter (Eds.), *AILA Review*, 8, 45-60.
- Brown, H. D. (2007). *Teaching by principles. An interactive approach to language pedagogy*. (3rd ed.). Addison-Wesley Longman Inc., White Plains, NY.
- Brumfit, C., & Johnson, K. (1979). *The Communicative Approach to Language Teaching*. Oxford: Oxford University Press.
- Carrell, P., & Eisterhold, J. (1983). Schema theory and ESL writing. *TESOL Quarterly*, 17(4), 553-573.
- Carell, P. L. (1983). Background knowledge in second language comprehension. *Language Learning and Communication*, 2(1), 25-33.

- Chall, J. S. (1987). Two vocabularies for reading: Recognition and meaning. In M. G. McKeown & M. E. Curtis (Eds.), *The Nature of Vocabulary Acquisition*. Hillsdale, NJ: Erlbaum.
- Clarke, M.A. (1980). The short-circuit hypothesis of ESL reading-or when language competence interferes with reading performance. *The Modern Language Journal*, 64, 203-209.
- Coady , J. (1997). L2 vocabulary acquisition through extensive reading. In J.Coady & T.Huckin (Eds.), *Second Language Vocabulary Acquisition*,(p. 225-237).
- Cooper, P. L. (1984). *The Assessment of Writing Ability: A Review of Research*. Princeton, NJ: Educational Testing Services.
- Cronbach, L. J. (1942). An analysis of techniques for diagnostic vocabulary testing. *Journal of Educational Research*, 3, 206-217.
- Day, R., & Bamford, J. (1998). *Extensive Reading in the Second Language Classroom*. Cambridge: Cambridge University Press.
- Durán, E. P. (1999). Teaching writing through reading: a text-centered approach. Retrieved May 30, 2008, from <http://www.aelfe.org/documents/text1-Duran.pdf>.
- Ellis, R. (1994). *The Study of Second Language Acquisition*. Oxford: Oxford University Press.
- Eskey, D. E. (1988). Holding in the bottom: An interactive approach to the language problems of second language readers. In P. Carrell, J. Devine, & D. Eskey (Eds.), *Interactive Approaches to Second Language Reading*. Cambridge, UK:Cambridge UP.
- Feng, A., & Byram, M. (2002). Authenticity in College English Textbooks - an Intercultural Perspective. *RELC Journal* 33, 58-84.

- Ferguson, Ch. (1959). Diglossia . *Word Journal* 15, 325-337.
- Gascoigne, C. (2005). Toward an understanding of the relationship between L2 reading comprehension and grammatical competence. *The Reading Matrix*, 5 (2).
- Goodman, K.S. (1968). *The Psychological Nature of the Reading Process*. Detroit: Wayne State University Press.
- Gough, P. B. (1972). One second of reading. In J.F. Kavanagh, & I.G. Mattingly (Eds.), *Language by Ear and Eye*. Cambridge: Cambridge University Press.
- Grabe, W. (1991). Current developments in second language reading research. *TESOL Quarterly*, 25, 375-406.
- Graves, M. F., Ryder, R., Slater, W., & Calfee, R. (2001). The Relationship between word frequency and reading vocabulary using six metrics of frequency. *Journal of Educational Research*, 81(2), 81-90.
- Haas, C., & Flower, L. (1988). Rhetorical Reading Strategies and the Construction of Meaning. *College Composition and Communication*, 39(2), 167-183.
- Henriksen, B. (1999). Three dimensions of vocabulary knowledge. *Studies in Second Language Acquisition*, 21, 303 -317.
- Hirsh, D. & P. Nation. 1992. What vocabulary size is needed to read unsimplified texts for pleasure?. *Reading in a Foreign Language* 8 (2), 689-696.
- Huang, C. (1999). *The effects of vocabulary knowledge and prior knowledge on reading comprehension of EFL students in Taiwan*. Unpublished Dissertation, Ohio University.
- Hu, M., & Nation, I. S. P. (2000). Vocabulary Density and Reading comprehension. *Reading in a Foreign Language*, 13(1), 403–430.

- Hymes, D. (1971). *On Communicative Competence*. Philadelphia: University of Pennsylvania Press.
- Kennedy, C., & Bolitho, R. (1985). *English for Specific Purposes*. Hong Kong: Macmillan.
- Khaldieh, S. A. (2001). The Relationship between Knowledge of Icrab, Lexical Knowledge, and Reading Comprehension of Nonnative Readers of Arabic. *The Modern Language Journal*, Volume 85 (3), 416-431.
- Koda, K. (1989). The effects of transferred vocabulary knowledge on the development of L2 reading proficiency. *Foreign Language Annals*, 22, 529-540.
- Krashen, S. (1981). *Second Language Acquisition and Second Language Learning*. New York: Pergamon.
- Laufer, B. & Sim, D. (1985). Measuring and explaining the reading threshold needed for English for academic purposes texts. *Foreign Language Annals*. 18(5), 405-411.
- Laufer, B. (1989). What percentage of text lexis is essential for comprehension? In Ch. Lauren and M. Nordman (Eds), *In Special Language: From Humans Thinking To Thinking Machine* (p. 316-323). Multilingual Matters.
- Laufer, B. (1992). How much lexis is necessary for reading comprehension? *In Vocabulary and Applied Linguistics*,(Eds), H. Bejoint and P. Arnaud. Macmillan, (p.126-132).
- Laufer, B. (1996). The lexical threshold of L2 reading: where it is and how it relates to L1 reading ability. In K Sajavaara and C. Fairweather,(Eds.), *Approaches to Second Language Acquisition* (p.55-62). Jyvaskyla: Cross Language Studies.
- Laufer, B. (1997). The lexical plight in second language reading: words you don't know, words you think you know and words you can't guess. In J. Coady and T.

- Huckin, (Eds.), *Second Language Vocabulary Acquisition: a Rationale for Pedagogy* (p. 20-34). Cambridge: Cambridge University Press.
- Lee, W. & D. L. Schallert. (1997). The relative contribution of L2 language proficiency and L1 reading ability to L2 reading performance: A test of the threshold hypothesis in an EFL context. *TESOL Quarterly*, 31, 713-739.
- Lee, W. (1995). Authenticity revisited: Text authenticity and learner authenticity. *ELT Journal* 49(4), 323–328.
- LeLoup, J. W. & Ponterio, R. (2007). Listening: You've got to be carefully taught. *Language Learning & Technology* 11(1), 4–15.
- Little, D. (1997). Responding authentically to authentic texts: a problem for self access learning? In: P. Benson and P. Voller, (Eds), *Autonomy and Independence in Language Learning* (p. 22-36). London: Longman.
- Little, D., Devitt, S., & D. Singleton (1989). *Learning foreign languages from authentic texts: Theory and practice*. Dublin: Authentic in Association with CILT.
- Liu, N. & Nation, I.S.P. (1985). Factors affecting guessing vocabulary in context. *RELC Journal* 16(1), 33-42.
- Martinez-Lang, A. (1995). Benefits of Keeping a Reading Journal in the Development of Second Language Reading Ability. *Dimension*, 65-79.
- Mishan, F. (2005). *Designing Authenticity into Language Learning Materials*. Bristol: Intellect.
- Mulling, S. (1991). The L2 Text: Genuineness and Authenticity. Retrieved June 30, 2008, from http://eric.ed.gov/ERICWebPortal/custom/portlets/recordDetails/detailmini.jsp?_nfpb=true&_ERICExtSearch_SearchValue_0=ED329119&ERICExtSearch_SearchType_0=no&accno=ED329119

- Nation, I.S. P. (1990). *Teaching and Learning Vocabulary*. New York: Newbury House.
- Nation, I. S. P. (2001). *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press.
- Nuttall, C. (1996). *Teaching Reading Skills in a Foreign Language*. Oxford: Heinemann.
- O'Malley, J. M., & Pierce, L. V. (1996). *Authentic Assessment for English Language Learners*. Boston, MA: Addison-Wesley Publishing Company.
- Peacock, M., (1997). The effect of authentic materials on the motivation of EFL Learners. *ELT Journal*, 51(2), 144 - 156.
- Popham, W. J. (1990). *Modern Educational Measurement* (2nd ed.). Boston: Allyn and Bacon.
- Qian, D. D. (2002). Investigating the relationship between vocabulary knowledge and academic reading performance: *An Assessment Perspective*. *Language Learning*, 52(3), 513-536.
- Richards, J. C. (1976). The role of vocabulary teaching. *TESOL Quartely*,10, 77-89.
- Ridgway, T. (1997). Thresholds of the background knowledge effect in foreign language reading, *Reading in a Foreign Language*, 11(1), 151-168.
- Rivers, W.M. (1981). *Teaching Foreign Language Skills* (2nd ed.). Chicago: University of Chicago Press.
- Ruddell, R. B., & Unrau, N.J. (1994). Reading as a meaning-construction process: The reader, the text, and the teacher. In R. B. Ruddell, M. R. Ruddell & H. Singer, *Theoretical Models and Processes of Reading*, (p. 996-1056). Newark, DE: International Reading Association.
- Rumelhart, D.E. (1980). Schemata: The Building Blocks of Cognition. In R.J. Spiro, B.C. Bruce, & W.F. Brewer (Eds.), *Theoretical Issues in Reading Comprehension*:

- Perspectives from Cognitive Psychology, Linguistics, Artificial Intelligence, and Education*. Hillsdale, NJ: Lawrence Erlbaum.
- Schmitt, N. and Meara, P. (1997). Researching vocabulary through a word knowledge framework: word associations and verbal suffixes. *Studies in Second Language Acquisition* 19, 17-36.
- Stahl, S. A. (1990). Beyond the instrumentalist hypothesis: some relations between word meanings and comprehension. *Technical Report No.505 of the Center for the Study of Reading*, University of Illinois at Urbana- Champaign.
- Singer, Harry, & Robert B. Ruddell. (1985). *Theoretical Models and the Processes of Reading*. (3rd ed.). Newark, DE: International Reading Association, Location: Dallas SIL Library 372.4 T396t. Interest level: specialist.
- Sweet, H. (1899). *The Practical Study of Languages*. London: J. M. Dent and Co.
- Swaffar, J. (1981). Reading in the foreign language classroom: Focus on process. *Die Unterrichtspraxis*, 14, 176-194.
- Swaffar, J. (1985). Reading authentic texts in a foreign language: A cognitive model. *The Modern Language Journal*, 69(1), 15-34.
- Van Lier, L. (1996). *Interaction in the Language Curriculum: Awareness, Autonomy, and Authenticity*. London: Longman.
- Vygotsky, L. (1986). *Thought and Language*. Cambridge, MA: The MIT Press.
- Wegener, D. T. & Petty, R. E. (1994). Mood management across affective states: The hedonic contingency hypothesis. *Journal of Personality & Social Psychology*, 66, 1034-1048.
- Widdowson, H. G. (1978). *Teaching Language as Communication*. Oxford: Oxford University Press.

- Yorio, C. (1971). Some sources of reading problems for foreign language learners. *Language Learning*, 21, 107-115.
- Young, D. J. (1993). Processing strategies of foreign language readers: Authentic and Edited Input. *Foreign Language Annals*, 26 (4), 451-468.

Appendix A. Background Questionnaire

Please write here last four digits of your BYU ID:.....

This questionnaire is part of a larger study, which you are encouraged but not obliged, to participate in. All the data and results of this study will be anonymous and will in no way affect your academic standing or your semester grades.

Demographic data

1. Age:-----
2. Gender : male female
3. Arabic course in which you are currently enrolled:.....
4. How long have you been studying Arabic?
5. Native language:-----
6. 5. Have you studied another foreign language? Yes() No(), if no , skip to question 8
7. If Yes, Please describe in the space provided below

Language	Number of years/where studied

Appendix B. Reading Comprehension Test - Passage 1.

Please write here last four digits of your BYU ID:.....

This test is part of a larger study, which you are encouraged but not obliged, to participate in. All the data and results of this study will be anonymous and will in no way affect your academic standing or your semester grades. The score of this test will remain confidential, and will only be used for research purposes. If you have any question, you may contact Shereen Salah at shereens@byu.edu.

Read the following instructions, before you take this test:

Please read the following passage thoroughly (you can read the passage as many times as you want, **(but no Dictionary use is allowed)**); then answer the questions below, **only in English**. Please make sure to use complete sentences in your answers.

ولد الدكتور أحمد زويل في 26 فبراير سنة 1946م في مدينة دمنهور بمصر. تلقى تعليمه الأولي في نفس المدينة ثم إنتقل مع الأسرة إلى مدينة دسوق حيث أتم تعليمه حتى المرحلة الثانوية. في سنة 1963م التحق أحمد زويل بكلية العلوم بجامعة الإسكندرية و حصل على بكالوريوس العلوم من قسم الكيمياء سنة 1967م، ثم نال بعد ذلك شهادة الماجستير من نفس الجامعة.

عمل زويل كمتدرب في شركة "شل" في الإسكندرية وأكمل دراساته العليا بعد ذلك في الولايات المتحدة الأمريكية، حيث نال درجة الدكتوراه من جامعة بنسلفانيا. بعد ذلك، إنتقل الدكتور زويل إلى جامعة بركلي بولاية كاليفورنيا وانضم لفريق الأبحاث هناك. وفي سنة 1976م عُين زويل في معهد كاليفورنيا التقني كأستاذ مساعد في الكيمياء الفيزيائية. في سنة 1982م تولى الدكتور زويل منصب أستاذ أول للكيمياء في معهد لينوس باولينج. و من أهم منجزاته هو أنه أصبح عضواً في الأكاديمية الأمريكية للعلوم في سن الثالثة و الأربعين، علماً أن هذه الأكاديمية لا تقبل أي عالماً بل تقبل أذكى العلماء و شرط أن يتخطى عمرهم الخامسة والخمسين عاماً.

Read the passage again thoroughly, and circle the words that you couldn't understand in the passage

ولد الدكتور أحمد زويل في 26 فبراير سنة 1946م في مدينة دمنهور بمصر. تلقى تعليمه الأولي في نفس المدينة ثم إنتقل مع الأسرة إلى مدينة دسوق حيث أتم تعليمه حتى المرحلة الثانوية. في سنة 1963م إلتحق أحمد زويل بكلية العلوم بجامعة الإسكندرية و حصل على بكالوريوس العلوم من قسم الكيمياء سنة 1967م، ثم نال بعد ذلك شهادة الماجيستر من نفس الجامعة.

عمل زويل كمتدرب في شركة "شل" في الإسكندرية وأكمل دراساته العليا بعد ذلك في الولايات المتحدة الأمريكية، حيث نال درجة الدكتوراه من جامعة بنسلفانيا. بعد ذلك، إنتقل الدكتور زويل إلى جامعة بركلي بولاية كاليفورنيا وانضم لفريق الأبحاث هناك. وفي سنة 1976م عُين زويل في معهد كاليفورنيا التقني كأستاذ مساعد في الكيمياء الفيزيائية. في سنة 1982م تولى الدكتور زويل منصب أستاذ أول للكيمياء في معهد لينوس باولينج.

و من أهم منجزاته هو أنه أصبح عضواً في الأكاديمية الأمريكية للعلوم في سن الثلاثة و الأربعين، علماً أن هذه الأكاديمية لا تقبل أي عالماً بل تقبل أنكى العلماء و شرط أن يتخطى عمرهم الخامسة و الخمسين عاماً.

Appendix C. Reading Comprehension Test – Passage 2.

Please write here last four digits of your BYU ID:.....

This test is part of a larger study, which you are encouraged but not obliged, to participate in. All the data and results of this study will be anonymous and will in no way affect your academic standing or your semester grades. The score of this test will remain confidential, and will only be used for research purposes. If you have any question, you may contact Shereen Salah at shereens@byu.edu.

Read the following instructions, before you take this test:

Please read the following passage thoroughly (you can read the passage as many times as you want, **(but no Dictionary use is allowed)**); then answer the questions below, **only in English**. Please make sure to use complete sentences in your answers.

إسرائيل تمنع المسلمين دخول باحة المسجد الأقصى- الجزيرة نت : 2007/2/4 م

منعت إسرائيل اليوم الأحد وصول المسلمين إلى باحة المسجد الأقصى تحسبا لاندلاع مظاهرات للاحتجاج على الأشغال التي يقوم بها الإسرائيليون قرب المسجد.

وقال المتحدث باسم الشرطة الإسرائيلية ميكي روزنفيلد إنه يمنع على الرجال والشباب المسلمين ممن تقل أعمارهم عن 45 عاما الوصول إلى باحة المسجد. ولم يشمل المنع النساء المسلمات، أما الرجال ممن تفوق أعمارهم 45 سنة فلن يدخل منهم إلا من يحمل بطاقة هوية زرقاء تسلمها إسرائيل.

وتأتي هذه الترتيبات بعدما دعا قاضي قضاة فلسطين ورئيس المجلس الأعلى للقضاء الشرعي تيسير التميمي الفلسطينيين للتوجه اليوم الأحد "إلى المسجد الأقصى المبارك للدفاع عنه وحمايته من المخططات والأخطار الإسرائيلية المحدقة به". ووجه التميمي نداءه بعد أن قررت إسرائيل هدم الطريق الذي يربط الحرم القدسي بباب المغاربة، وهو ما يعد مساسا بأجزاء من الحرم القدسي الشريف.

Questions: (Answer in **English only**)

- 1- What did the Israelis do on Sunday?
- 2- What is the reason behind that?
- 3- Who will be affected by Israel's decision mentioned in the first paragraph?
- 4- According to the last paragraph, what has Israel decided to do?
- 5- What do the Palestinians want to do, in response to Israel's decision, according to the last passage?

Read the passage again thoroughly, and circle the words that you couldn't understand in the passage

إسرائيل تمنع المسلمين دخول باحة المسجد الأقصى - الجزيرة نت 2007/2/4 م

منعت إسرائيل اليوم الأحد وصول المسلمين إلى باحة المسجد الأقصى تحسبا لاندلاع مظاهرات للاحتجاج على الأشغال التي يقوم بها الإسرائيليون قرب المسجد.

وقال المتحدث باسم الشرطة الإسرائيلية ميكي روزنفيلد إنه يمنع على الرجال والشباب المسلمين ممن تقل أعمارهم عن 45 عاما الوصول إلى باحة المسجد. ولم يشمل المنع النساء المسلمات، أما الرجال ممن تفوق أعمارهم 45 سنة فلن يدخل منهم إلا من يحمل بطاقة هوية زرقاء تسلمها إسرائيل.

وتأتي هذه الترتيبات بعدما دعا قاضي قضاة فلسطين ورئيس المجلس الأعلى للقضاء الشرعي تيسير التميمي الفلسطينيين للتوجه اليوم الأحد "إلى المسجد الأقصى المبارك للدفاع عنه وحمايته من المخططات والأخطار الإسرائيلية المحدقة به". ووجه التميمي نداءه بعد أن قررت إسرائيل هدم الطريق الذي يربط الحرم القدسي بباب المغاربة، وهو ما يعد مساسا بأجزاء من الحرم القدسي الشريف.

Appendix D. Lexical Coverage Test – Passage 1.

Please write here last four digits of your BYU ID:.....

Read the passage thoroughly, and translate the words listed below or paraphrase their meanings based on the given text. (No Dictionary use is allowed)

ولد الدكتور أحمد زويل في 26 فبراير سنة 1946م في مدينة دمنهور بمصر. تلقى تعليمه الأول في نفس المدينة ثم إنتقل مع الأسرة إلى مدينة دسوق حيث أتم تعليمه حتى المرحلة الثانوية. في سنة 1963م إلتحق أحمد زويل بكلية العلوم بجامعة الإسكندرية و حصل على بكالوريوس العلوم من قسم الكيمياء سنة 1967م، ثم نال بعد ذلك شهادة الماجيستر من نفس الجامعة.

عمل زويل كمتدرب في شركة "شل" في الإسكندرية وأكمل دراساته العليا بعد ذلك في الولايات المتحدة الأمريكية، حيث نال درجة الدكتوراه من جامعة بنسلفانيا. بعد ذلك، إنتقل الدكتور زويل إلى جامعة بركلي بولاية كاليفورنيا وانضم لفريق الأبحاث هناك. وفي سنة 1976م عُين زويل في معهد كاليفورنيا التقني كأستاذ مساعد في الكيمياء الفيزيائية. في سنة 1982م تولى الدكتور زويل منصب أستاذ أول للكيمياء في معهد لينوس باولينج.

و من أهم منجزاته هو أنه أصبح عضواً في الأكاديمية الأمريكية للعلوم في سن الثلاثة و الأربعين، علماً أن هذه الأكاديمية لا تقبل أي عالماً بل تقبل أذكى العلماء و شرط أن يتخطى عمرهم الخامسة والخمسين عاماً.

أتم/يتم

أنتقل/ ينتقل

تلقى/ يتلقى

المرحلة

حتى

تعليمه (تعليم)

التحق / يلتحق	حصل على / يحصل على	كلية
قسم	الكيمياء	شهادة
نال / ينال	عمل / يعمل	متدرب (يتدرب)
شركة	أكمل / يكمل	درجة
أنضم / ينضم	فريق	الابحاث
عين	معهد	التقني
تولى / يتولى	منصب	لينوس باولينج
منجزاته (إنجازات)	أهم	عضواً
سن	علماً	قبلت / تقبل
عالمأ	أذكى	شرط
يتخطى / تخطى	عمرهم (عمر)	عامأ

Appendix E. Lexical Coverage Test - Passage 2.

Please write here last four digits of your BYU ID:.....

Read the passage thoroughly, and translate the words listed below or paraphrase their meanings based on the given text. (No Dictionary use is allowed)

منعت إسرائيل اليوم الأحد وصول المسلمين إلى باحة المسجد الأقصى تحسباً لاندلاع مظاهرات للاحتجاج على الأشغال التي يقوم بها الإسرائيليون قرب المسجد.

وقال المتحدث باسم الشرطة الإسرائيلية ميكي روزنفيلد إنه يمنع على الرجال والشباب المسلمين ممن تقل أعمارهم عن 45 عاماً الوصول إلى باحة المسجد. ولم يشمل المنع النساء المسلمات، أما الرجال ممن تفوق أعمارهم 45 سنة فلن يدخل منهم إلا من يحمل بطاقة هوية زرقاء تسلمها إسرائيل.

وتأتي هذه الترتيبات بعدما دعا قاضي قضاة فلسطين ورئيس المجلس الأعلى للقضاء الشرعي تيسير التميمي الفلسطينيين للتوجه اليوم الأحد "إلى المسجد الأقصى المبارك للدفاع عنه وحمايته من المخططات والأخطار الإسرائيلية المحدقة به". ووجه التميمي نداءه بعد أن قررت إسرائيل هدم الطريق الذي يربط الحرم القدسي بباب المغاربة، وهو ما يعد مساساً بأجزاء من الحرم القدسي الشريف.

تحسباً

باحة

وصول (يصل)

منعت/ يمنع

الأشغال

إحتجاج (يحتج)

مظاهرات

إندلاع

المتحدث	الشباب	ممن	تقل
أعمار (عمر)	لم	شمل / يشمل	تفوق
حمل / يحمل	بطاقة هوية	زرقاء	تأتي
الترتيبات	بعدها	دعا / يدعو	قاضي
القضاء	التوجه	الدفاع (يدافع)	حمايته (حماية)
المخططات (يخطط)	الأخطار	المحدقة به	نداءه (ينادي)
قررت / تقرر	هدم (يهدم)	ربط / يربط	الطريق
يُعد	مساساً	الحرم القدسي الشريف	أجزاء

باب المغاربة