The American University in Cairo
School of Humanities and Social Sciences

Investigating the Effects of Using Concordance Data on Vocabulary Acquisition in
an Egyptian English for Academic Purposes Setting

A thesis submitted to

The English Language Institute
Department of Teaching English as a Foreign Language

in partial fulfillment of the requirements for
the degree of Master of Arts

by

Jenna M. Steiner

May 2011
The American University in Cairo

Investigating the Effects of Using Concordance Data on Vocabulary Acquisition in
an Egyptian English for Academic Purposes Setting

A thesis submitted by Jenna M. Steiner
to the Department of Teaching English as a Foreign Language
English Language Institute
May 2011
in partial fulfillment of the requirements for
the degree of Master of Arts
has been approved by

Dr. Amira Agameya
Thesis Committee Chair _________________________

Dr. Lori Fredricks
Thesis Committee Second Reader _________________________

Dr. Atta Gebril
Thesis Committee Third Reader _________________________

_________________________________  _________  __________________________  _________
Department Chair                   Date                 Dean                      Date
To my parents

Charles and Lori Steiner

for their endless encouragement

and unwavering support of me in all my endeavors

“If you raise your children to feel that they can accomplish any goal or task they decide upon, you will have succeeded as a parent and you will have given your children the greatest of all blessings.”

Brian Tracy
Acknowledgements

First and foremost, I would like to thank my first reader, Dr. Amira Agameya, whose Corpus Linguistics class sparked my interest in this topic. I owe her my deepest gratitude for all her encouragement and support, without which this thesis would not have been possible. Thank you, Dr. Agameya.

I would also like to thank my second reader, Dr. Lori Fredricks, for all her valuable input throughout this process and my third reader, Dr. Atta Gebril, who always made his support available to me and provided me with extensive feedback. Thank you to both of you for your patience and guidance.

Finally, I would like to thank the six English Language Institute teachers who so graciously allowed me to conduct my research in their classrooms. I give my sincerest thanks to all of you for your patience and willingness to help.
Abstract

The purpose of the present study was to investigate the effects of using concordance data in the classroom on vocabulary acquisition in the English Language Institute (ELI) at the American University in Cairo (AUC). Specifically, the effects on overall lexical knowledge and ability to distinguish between synonyms were explored in addition to students’ attitudes towards the use of concordance data.

This study employed an exploratory and quantitative approach and used a convenience sample of six intact Egyptian EFL university classes. Three classes with a total of 26 participants were randomly assigned to the experimental group and the remaining three classes with a total of 24 participants were assigned to the control group. The researcher administered a pretest, a vocabulary lesson, and a posttest to each group. The experimental group also completed a questionnaire on their feelings towards the use of concordance lines.

The results of the quantitative data indicated that both the experimental and the control group made statistically significant gains from pretest to posttest. However, neither group made higher gains than the other, thus suggesting that there was no difference between the two groups in overall lexical knowledge or ability to distinguish between synonyms. Meanwhile, the qualitative data revealed a positive attitude of behalf of the participants towards the use of concordance lines in their classrooms.
Table of Contents

Chapter I: Introduction ........................................................................................................1
  1.1 Statement of the Problem..............................................................1
  1.2 Research Questions......................................................................2
  1.3 Definitions of Terms.................................................................3
  1.4 Assumptions, Delimitations, and Limitations..............................4
  1.5 Significance of Study...............................................................5
  1.6 Summary of the Problem.........................................................5

Chapter II: Literature Review..................................................................................7
  2.1 Introduction to Literature Review ...........................................7
  2.2 Academic Word List (AWL)......................................................8
  2.3 Recent Research on Vocabulary Acquisition..........................8
  2.4 Recent Research on Classroom use of Corpora......................12
  2.5 Relation of Research to Present Study..................................17

Chapter III: Methodology ..............................................................................19
  3.1 Participants...............................................................................19
  3.2 Sampling Procedures.............................................................21
  3.3 Materials and Instruments.....................................................21
     3.3.1 Materials...........................................................................21
     3.3.2 Instruments......................................................................22
  3.4 Data Collection Procedures..................................................25
  3.5 Data Analysis...........................................................................27

Chapter IV: Results...................................................................................30
  4.1 Results of Vocabulary Test Data...........................................30
  4.2 Results of Questionnaire Data.................................................34

Chapter V: Discussion and Conclusion....................................................39
  5.1 Discussion ..............................................................................39
  5.2 First Research Question.........................................................39
  5.3 Second Research Question....................................................40
  5.4 Third Research Question.......................................................40
List of Tables

Table 3.1 TOEFL Placement Scores for English 99 and English 100..........................20
Table 3.2 Test Specifications.................................................................................25
Table 4.1 Descriptive Statistics.............................................................................31
Table 4.2 Multivariate Test.....................................................................................32
Table 4.3 Tests of Within-Subjects Effects.............................................................33
Table 4.4 Tests of Between-Subjects Effects..........................................................34
Table 4.5 Question 1: What do you like about using the concordance lines to complete the in class activity?.................................................................35
Table 4.6 Question 2: What did you dislike about using the concordance lines to complete the in class activity?.................................................................36
Table 4.7 Question 3: What kind of information did you learn about each word based on the concordance lines?.................................................................37
Table 4.8 Question 4: Do you think the concordance lines and in class activity helped you to learn the vocabulary words? If no, why not?...............................37
Table 4.9 Question 5: Would you want your teacher to use concordance lines in class again? Why or why not?.................................................................38
Chapter I

Introduction

The first introduction to the field of corpus linguistics came in 1967 when Henry Kucera and Nelson Francis published their article “Computational analysis of present-day American English” based on the Brown Corpus. The early uses of corpora focused on analyses of the English language. Corpora were not readily available for public use until the 1990’s when there was an explosion of corpora including the British National Corpus and the International Corpus of English. Side-by-side with the development of these corpora came the development of concordancing software which made sorting through and analyzing these corpora easy and efficient. While corpora and concordancing software have been available since the early 1990’s, the presence of computers both in schools and homes did not become prevalent until much later. Now that computers have become ubiquitous through many parts of the world, it is an ideal time to investigate the effects that classroom use of corpus data may have on vocabulary acquisition in an Egyptian English for academic purposes (EAP) setting (Meyer, 2002).

The present study focuses on the effects of teaching with concordance data on vocabulary acquisition, as opposed to grammar or any other language skill, because vocabulary knowledge is a foundation upon which other language skills, specifically reading, are built (Jun Zhang, 2008). As Cobb (2007) says, this revelation of vocabulary knowledge as a foundational language skill has revolutionized the area of vocabulary acquisition. Nowadays, researchers have a strong interest in the study of vocabulary acquisition, as we can see by the recent explosion of studies related to the area, and researchers and teachers alike seem to be scrambling to find the best way to teach vocabulary.
1.1 Statement of the Problem

Literature focusing on the area of vocabulary acquisition has moved away from the use of word lists in teaching vocabulary items towards methods that involve teaching words in context (Collins, 2010; Nation, 2001; Verhallen & Bus, 2010; Walters & Bozkurt, 2009). Since corpora allow students to see an abundance of contexts for any given vocabulary item in addition to facilitating the implicit learning of these items, it seems logical that classroom use of concordance data could have positive effects on vocabulary acquisition.

Studies such as Kaur and Hegelheimer (2005), Liu and Jiang (2010), and Varley (2009) look at the effects of classroom corpus use and all encounter similar problems. Students, in general, find navigating corpora difficult and overwhelming and are hesitant to use them when an alternative is available. Therefore, a study is needed that tests the effects of classroom corpus use on vocabulary acquisition while also finding a way to ease students’ hesitance and dislike towards the use of corpora.

This study investigates the effects of using concordance data as a teaching tool on vocabulary acquisition in the English Language Institute (ELI) at the American University in Cairo (AUC). The ELI is an EAP program designed to improve the English proficiency level of AUC students whose TOEFL scores did not meet the required cut-off score for full matriculation into the university. The TOEFL scores of the participants are presented in table 3.1. This study also explores students’ attitudes towards the use of concordance data in their classroom.
1.2 Research Questions

The present study is designed to provide answers to the following questions:

1.) Does classroom use of concordance data have an effect on vocabulary acquisition by advanced EFL university students?

2.) Does classroom use of concordance data lead to a better distinction between near-synonyms for these advanced EFL university students?

3.) What are the students’ attitudes towards the use of concordance data in their classroom?

1.3 Definition of Terms

*Vocabulary acquisition:* This term is used when an individual gains knowledge of a previously unknown word. This knowledge can be either receptive, productive or a mixture of both. It is important to note that not all vocabulary knowledge can be classified as strictly receptive or strictly productive. It is best to think of vocabulary knowledge as more of a continuum. For the purposes of this study, vocabulary acquisition is measured using a “recall aided by recognition” test which requires participants to have more than a receptive knowledge of vocabulary items but does not require them to be able to produce these items freely in speech or writing (Kelley & Krey, 1934).

*Corpus/corpora:* A database or collection of authentic utterances of a given language.

*Corpus use:* The use of an activity that requires students to look at concordance lines and other data obtained from the corpus such as frequent collocates and phraseology.

*Concordance line:* A single example extracted from a corpus.
**Concordance data:** The students themselves do not use the concordancing software, but rather the researcher, or teacher, makes decisions about what concordance lines to select, how many to include and what order to put them in. Students are then given a handout with the concordance lines.

**Distinction:** For the purposes of this study, distinction is used to mean that one word is not misused as its partner in the near-synonym word pairs.

**Synonyms:** For the purposes of this article, “synonym” in fact refers to “near-synonyms” as defined by Inkpen & Hirst (2006) and Inkpen (2007). “Near-synonyms” are words that are almost synonyms but are not “absolute synonyms” because they are not completely inter-substitutable. They may have different collocations, connotations, or implications.

**1.4 Assumptions, Delimitations, and Limitations**

This study is operating under the assumption that vocabulary acquisition can be measured through the use of a pretest and posttest modeled after the productive section of Paul Nation’s Vocabulary Levels Test found in *Learning Vocabulary in another Language* (Nation, 2001). It also assumes that the answers given by the participants on the questionnaire are truthful.

As for delimitations, this study aims to explore the effects of classroom use of concordance data on vocabulary acquisition only. Therefore, for the purposes of this study, effects on grammar, reading, and writing are not investigated. In addition to this, free productive knowledge of the vocabulary items is not investigated either. The study also uses pre-selected concordance data, as defined above, as a way to ease the negative
attitude students have towards corpora as reported in Kaur and Hegelheimer (2005), Liu and Jiang (2010) and Varley (2009).

Due to the small and specific nature of the sample size, the major limitation of this study is that it cannot be confidently generalized to a larger population. Also, due to the limited access to classes in the ELI, the duration of the treatment is only one hour which means the results are not necessarily representative of those that would be obtained from a significantly longer treatment. Therefore, further studies with larger sample sizes and longer treatments are needed to verify the results of this study.

1.5 Significance of Present Study

This study is important because it is investigating a unique way of using corpus data inside the classroom. This method of using pre-selected concordance lines aims to change students’ negative attitudes towards working with corpora (Kaur & Hegelheimer, 2005; Liu & Jiang, 2010; Varley, 2009). In addition, the results show whether or not classroom use of concordance data help facilitate vocabulary acquisition specifically in the distinction between near-synonyms, which is a problematic area for non-native speakers of English (Inkpen & Hirst, 2006). While the results of this study may not necessarily be generalizable outside of the ELI at the AUC context, they pave the way for further research on this topic where the study can be replicated with a larger sample size and the results verified.

1.6 Summary of the Problem

In summary, the study investigates the effects, if any, of using concordance data in the classroom on the vocabulary acquisition of ELI students at the AUC. In particular, the knowledge of the target vocabulary items, as defined above, and the ability to
distinguish between near-synonyms are measured. The attitudes of students towards the use of the concordance data are also addressed. The following section shows how this study fills a gap in the literature relating to the use of corpus data in teaching vocabulary.
Chapter II

Literature Review

2.1 Introduction to Literature Review

The literature review of this study focuses on the literature illustrating the importance of the Academic Word List (AWL), discussing the most recent trends in vocabulary acquisition, and dealing with classroom use of corpus data. The literature on the AWL shows the importance of vocabulary as a foundational skill and the need to focus on high frequency academic words in an EAP program. In addition, the vocabulary classes in the ELI teach from the AWL, and so this study uses the target words from the AWL. The literature on vocabulary acquisition calls for contextual teaching of vocabulary where the students have multiple exposures to the word, are encouraged to infer the meaning, and then complete a vocabulary-specific activity. The use of concordance data, as designed in this study, meets all these criteria. Finally, despite the fact that literature on student use of concordance data is limited (Breyer, 2009; Chambers, 2005), there continues to be evidence of the positive pedagogical implications of student corpus use and a call for further research and classroom implementation.

Much of the literature discussed in this literature review was found using an online database such as ERIC, Google Scholars, and the AUC’s Library Database. First, the terms “corpus” and “vocabulary acquisition” were searched for together. This resulted in very few results as there have not been many studies investigating the effects of corpus use on vocabulary acquisition. However, a very large number of results were found when the search term “vocabulary acquisition” was used. Therefore, the studies included in this literature review were mostly limited to those published in the last five
years since the focus is on current trends in vocabulary acquisition. Exceptions were made for some studies in order to show the progression in trends in vocabulary acquisition. Lastly, studies found using “corpus” as a search term were limited to those that investigated in-class use of corpus since that is the focus of this study. Other studies discussing the use of corpus for curriculum design and materials development were excluded.

2.2 Academic Word List

The AWL compiled by Averil Coxhead in 1998, is based on a corpus database of about 3.5 million words. The list is compiled from the academic written text genre of the corpus and looks at words outside of the 2,000 most frequent words in the English language. The AWL contains 570 word families which consist of the headword, or stem form of the word, in addition to all other inflectional and derivational forms of the word. There are approximately 3,000 words in total. While these words make up 10 percent of the academic genre of the corpus, they only make up about 1.4 percent of a similarly sized fiction corpus, thus showing that the words are academic in nature (Coxhead, 2002). At this time, the AWL is the predominant focus of vocabulary classes in the Intensive English Program (IEP) at the AUC. Also, the primary textbook for teaching vocabulary in the IEP, Focus on Vocabulary (Nation & Gu, 2007), focuses entirely on the AWL. For this reason, this study chose the target vocabulary words from the AWL in order to align the goals of this study with the goals of the teachers in the ELI.

2.3 Recent Research on Vocabulary Acquisition

In order to argue for the potential pedagogical implications of concordance data on vocabulary acquisition, it is important to have an idea of how students learn
vocabulary and the problems they encounter. There have been a plethora of studies published in just the past few years focusing on strategies to improve vocabulary acquisition. One prevalent strategy is to teach vocabulary through reading. Collins (2010) and Mason, Vanata, Jander, Borsch, and Krashen, (2009) investigated second language vocabulary acquisition through storybook reading. According to Collins (2010), using “rich explanation” of words helps facilitate vocabulary acquisition because it involves drawing attention to the vocabulary word as it is used in context and provides other examples in context. He concluded that this “rich explanation” (i.e. teaching the word in context) is a large contributing factor to vocabulary acquisition (Collins, 2010).

Meanwhile, Mason et al. (2009) compared one group of students who were read a story containing certain vocabulary items to another group who received those vocabulary items in list form (i.e. the vocabulary word side-by-side with its L1 equivalent). Mason et al. (2009) found that the students who were read the story, and were therefore taught the words in context, acquired nearly twice as many vocabulary items as the group who received the list of items and L1 equivalents. A study by Verhallen and Bus (2010) also found that teaching vocabulary items in context has a positive effect on vocabulary acquisition.

If, according to Verhallen and Bus (2010), Mason et al. (2009), and Collins (2010), learning words through reading is very effective one has to wonder why this is the case. In the studies above, it is hypothesized that reading allows for multiple exposures to target vocabulary items, and this may be a contributing factor in vocabulary acquisition. Zahar, Cobb and Spada (2001) investigate this idea of multiple exposures. In this study, exposure is used to describe how many times the students encounter a given
vocabulary item in a single reading. They argue that while reading does facilitate vocabulary acquisition, often times the words are not repeated enough times for true acquisition to take place.

Hulstijn, Hollander, and Greidanus (1996) also found that reading, in-and-of-itself, is not sufficient for vocabulary acquisition to take place. They conclude that the incidental learning of vocabulary that takes place through reading is responsible for only a small amount of total vocabulary knowledge. Again, this is because vocabulary items are not repeated frequently enough in a given reading. The use of concordance data may be able to answer this call for multiple exposures of vocabulary items. With concordance lines, students can quickly read through five or ten lines which expose them to the same vocabulary item multiple times in a variety of contexts and at an intensified rate.

Since reading alone is not sufficient for vocabulary acquisition to take place, studies such as Atay and Kurt (2006) call for post-reading tasks to solidify vocabulary acquisition. They investigate the effects of discrete written tasks and interactive tasks as post-reading activities. They found that both groups made significant gains in terms of knowledge of the target vocabulary and that there was no significant difference between the two groups. Therefore, they concluded that the emphasis on the target vocabulary and the repeated exposure of these items in a lesson are the crucial elements in explicit vocabulary instruction (Atay & Kurt, 2006).

In addition to Atay and Kurt (2006), Walters and Bozkurt (2009), and Zimmerman (1997) call for explicit vocabulary instruction. Walters and Bozkurt (2009), for example, investigate the effects of keeping vocabulary journals on vocabulary acquisition. In these journals, the students were asked to fill in information on each word
such as alternate meanings, parts of speech, synonyms, and antonyms and also to write down any time they read or hear the word in context. The experimental group scored significantly higher than the control group on both receptive and productive knowledge of vocabulary (Walters & Bozkurt, 2009).

Zimmerman (1997) also calls for the integration of reading and post-reading explicit vocabulary instruction as this explicit instruction allows for even more exposure to the vocabulary items. She found that students attending an intensive English program in preparation for university, very similar to the ELI at the AUC, had a significant gain in their common academic words used across disciplines as compared to students exposed to reading only. She concludes by advising teachers to teach vocabulary items in context, provide multiple exposures to the words in a session and to use a post-reading communicative vocabulary activity.

It can be concluded from the results of the studies discussed above that students learn vocabulary better when the words are presented in context (Mason et al., 2009; Zimmerman, 1997) and the students are exposed to the words multiple times in a session (Atay & Kurt, 2006; Nation, 2001; Zahar, 2001; Zimmerman, 1997). However, presenting the words in context without any sort of explicit vocabulary instruction is not sufficient (Brown, 1993), and so explicit vocabulary instruction is called for in language classrooms (Nation, 2001; Walters & Bozkurt, 2009; Zimmerman, 1997). Many studies argue that students learn vocabulary best when they are left to infer the meaning themselves based on context (Hulstijn, 1992). However, Grabe (2009) has shown that students often infer incorrectly which is problematic. Therefore, the need for an explicit
vocabulary task in addition to multiple contextual exposures to the vocabulary item is again stressed.

Since corpora are able to provide an abundance of contextual examples for any given vocabulary item and require the students to decipher meaning and use on their own, it can be hypothesized that the presentation of concordance data, when combined with an explicit vocabulary activity, will have promising implications on vocabulary acquisition. This is especially true when it comes to distinguishing near-synonyms since Inkpen (2007) concludes that context and collocations, both of which corpora provide, are the best way for EFL students to distinguish between near-synonyms.

2.4 Recent Research on Classroom Use of Corpora

The previous studies deal with vocabulary acquisition without the aid of corpora as it is important to get a sense of what is being studied and proposed in the area of vocabulary acquisition in order to discuss potential uses of concordance data in acquiring vocabulary. As for the studies dealing directly with student use of corpora, the number of is relatively limited (Breyer, 2009; Chambers, 2005). The overwhelming majority of studies dealing with corpora investigate how it can be used to develop textbooks (Wang & Good, 2007) and curriculum (Shirato & Stapleton, 2007) and to analyze how language is used and the subsequent implications for teaching (Biber, Conrad & Reppen, 1998). For the purposes of this review, however, these types of studies are not discussed as the focus of this study is the implementation of corpus use within the classroom as a teaching tool. The relatively limited number of these types of studies is yet even further evidence that research on this topic is very much needed.
The first recommendations for corpus use in the classroom came from Tim Johns (1986) when he published his paper on the data-driven learning approach. The data-driven learning approach called for the exclusive use of authentic materials and a focus on exploratory tasks, learner-centered activities, and student exploitation of tools. Johns (1986) proclaimed that student use of corpora was the way to achieve these goals. Unfortunately, despite the continued enthusiasm by researchers concerning corpus linguistics, research on corpus use in the classroom still has yet to become prevalent. In fact, Nation (2001) claims that as of the publishing of his book *Learning Vocabulary in another Language* in 2001, there was only one experimental study on student use of concordance lines which was Cobbs (1997).

Cobbs (1997) conducted a study to find out if there were any measurable effects on vocabulary acquisition from student use of concordancing. He introduced concordance lines based on the students’ weekly vocabulary lists and had them perform class activities where they chose the correct definition from multiple choices, filled-in-the-blank with the correct word, etc. Based on an analysis of a pretest and posttest, he found that students scored significantly higher when using concordance information than when not. Nation (2001) looked at this study and added that in-class use of concordance data promotes vocabulary learning by presenting authentic context and rich information about collocates, grammatical patterns, and word families. He emphasizes this idea of discovery learning as one of the best ways to learn vocabulary. While this study and its results were very promising, it was conducted over a decade ago, and therefore it is important to also look at more recent literature on the topic.
Varley (2009) is an important study that looks into students’ attitudes towards using concordancing software programs as opposed to more traditional resources such as dictionaries and grammar books. Varley (2009) investigates how the use of a corpus can complement traditional teaching methods, how students perceive concordance software and how this use of concordance data affects vocabulary acquisition. The participants were 19 EFL students in New Zealand mostly of Chinese decent. They had two hours of classroom-based learning and two hours of lab-based learning per week for 14 weeks. The tasks included creating frequency lists, coming up with new definitions, and discovering collocation information.

The participants took a questionnaire before the course, kept a reflective log during the 14 weeks, and participated in occasional discussions where they were asked for more detailed feedback. In the end, the majority of students either agreed or highly agreed that concordancing is helpful in language learning and most said they would probably use it again. Unfortunately, this study does not empirically show whether or not the concordancing software improved the language of the participants. Rather, it only investigated the students’ attitudes and perceived benefits of concordancing. In addition to this, it is important to note that despite the students’ claims that concordancing is, in fact, helpful when it comes to language acquisition, many strongly disliked using concordance software because they found it too “confusing” and “overwhelming” (Varley 2009).

From another perspective, Breyer (2009) investigated the implementation of corpus use in the classroom from the teacher’s perspective. She discusses in great detail the point that within the academic field there is much excitement about the potential of
corpus use, but unfortunately this has not yet transferred to current language teaching
practices. Her study looks at student teachers who were trained to use corpus and then
later implemented this corpus use in their classrooms.

Breyer found in her study that by the end of the study the student teachers had
come to recognize the value of concordancing as a tool for the learner to explore the
complexities of language. The student teachers also felt that corpus use lent credibility to
their lessons by allowing the learner to explore authentic texts and discover language use
at their own pace In the end, Breyer concludes that a lack of teacher training concerning
corpus use and a lack of readily available materials are the main reasons why corpus use
has yet to become prevalent in classrooms. Breyer (2009) calls for further research to
empirically test the effects of classroom corpus use in order to encourage teachers to
implement corpus use in their classrooms.

The final two studies discussed in this review are very important to this study
because they actually tested the use of concordance data in classrooms. First, Kaur and
Hegelheimer (2005) investigated whether or not the use of an online concordance
program would facilitate the transfer of academic words to a writing task. It was found
that while the experimental group did not perform better on either of the in-class
vocabulary activities, they did score higher on the writing task in terms of frequency of
use and accuracy.

This study hopes to verify these results while simultaneously addressing two
major limitations of the study. First and most importantly, Kaur and Hegelheimer (2005)
admitted that many participants in the experimental group did not make much use of the
concordance program since they were allowed to use the online dictionary which they
were more familiar with. If the participants hardly used the concordance program, then how could it have aided in the transfer of the target words? For this reason, the experimental group in this study was only allowed to use the concordance data provided. Secondly, the researchers only allowed 30 minutes for each vocabulary task, which, is not nearly enough time for someone to use a concordance program sufficiently. For this reason, in the interest of maximizing classroom time, this study provided students with a limited number of concordance lines. This is in place of the students having to work the concordancing software and sort through a plethora of concordance lines themselves, thus wasting precious class time.

The most recent study on classroom corpus use comes from Liu and Jiang (2010), who implemented a corpus-based lexico-grammatical approach to teaching grammar in three classrooms and studied the benefits of this approach. While this study investigated grammar instead of vocabulary it is still important because it investigated the effects of implementing a corpus-based approach in the classroom. Liu and Jiang (2010) claimed that the implementation of a corpus-based lexico-grammatical approach benefited students in that by the end of the semester they had a better command of grammatical rules/patterns, they better understood how great an effect context can have on meaning, and they had a better understanding of grammar.

Unfortunately, these conclusions relied only on the students’ own opinions of what they had learned. The researchers, in fact, very rarely referred to the student work they collected in order to empirically show that the students had greatly improved in their understanding of grammar. Even if they had proved that the students improved greatly, without a control group the researcher would never know for sure that it was the new
approach that had caused this improvement. The limitations in Liu and Jiang (2010) make it clear that quantitative studies are needed that empirically show the effects of corpus use on language learning which is what this study aims to do.

While the studies above show promising potential for classroom use of concordance data, many of the studies have critical weaknesses that must be addressed. For example, Kaur and Hegelheimer (2005) had to admit that the participants in their study did not make much use of the concordancing program, which calls their conclusions into question. How could the concordancing program be responsible for the vocabulary acquisition if the students themselves admitted they did not make use of it? Also, they allowed for only 30 minutes for the classroom activity which is not nearly enough time to sort through thousands of concordance lines on each vocabulary item. In addition to this, many of the studies mentioned did not empirically test the results of using concordance data (Liu & Jiang, 2010; Varley, 2009). For this reason, it is crucial that more studies be conducted that empirical test the effects of student use of concordance data much like Cobbs (1997), and this is what this study proposes to do.

2.5 Relation of Research to Present Study

Recent studies in the area of vocabulary acquisition call for a new method that replaces traditional memorization of vocabulary items with a method that provides plenty of exposure to these items in context and involves deeper thought processes on the part of the student (Mason et al., 2009). The use of corpus data in the classroom answers this call as it allows the students to see the vocabulary item in an abundance of contexts and forces them to decipher the proper form, meaning, and use on their own.
As for the studies relating to classroom use of corpus data, a common theme is one of student dislike towards the use of corpora (Haur & Hegelheimer, 2005; Liu & Jiang 2010; Varley, 2009). Therefore, it is the aim of this study to use teacher-selected concordance data as a way of easing the frustration expressed by students when working with corpus databases. In addition to this, there were critical weaknesses in many of these studies such as results not being empirically verifiable (Liu & Jiang, 2010) and students avoiding the use of corpus data if possible (Kaur & Hegelheimer, 2005), and this study aims to address these weaknesses.
Chapter III

Research Methodology

The present study was designed in order to investigate the effects of using concordance data in the classroom. Therefore, a pretest-posttest design was used with one experimental group and one control group. The study used an exploratory and mixed-method approach. The participants and sampling procedures of this study are discussed in the following two sections, followed by an explanation of the materials and instruments, data collection procedures, and finally data analysis.

3.1 Participants

The participants in this study were from six intact classes in the English Language Institute (ELI) at the AUC in the spring 2011 semester. There were 50 participants with a total of 24 participants in the experimental group and 26 in the control group. The AUC is an English-medium university meaning all classes, lectures, readings, and assignments are conducted in English. This leads to an interesting environment where students must interact in English with their classmates and professors, at least in class, even though this is not their native language. Students enrolled in classes in the ELI have been accepted into the university, but their TOEFL scores did not meet the required score for full matriculation. Therefore, the ELI offers EAP courses for those students in writing, reading, and vocabulary with the goal of raising their English proficiency level in preparation for classes at the AUC.

Of the six participating classes, two were at the English 99 level and four were at the English 100 level. All of the classes that volunteered to participate in the study were used as there were a relatively low number of students enrolled in the ELI this semester.
Traditionally, enrollment in the ELI is much lower in the spring semester than in the fall semester. Due to this fact, there were not enough students in either level to constitute a large enough sample size so two levels were used. There were a total of 18 participants from the English 99 level and 32 from the English 100 level, but this study did not look at proficiency level as a variable. The placement cut-off scores for both levels are provided in Table 3.1 in order to give an idea of the proficiency level of the participants.

Table 3.1

*Placement Cut-off Scores for English 99 and English 100*

<table>
<thead>
<tr>
<th>Placement level</th>
<th>iBT TOEFL Score</th>
<th>iBT Writing Score</th>
<th>PPT TOEFL Score</th>
<th>CBT TOEFL Score</th>
<th>TWE Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>English 99</td>
<td>62-75</td>
<td>17-19</td>
<td>503-539</td>
<td>177-204</td>
<td>3.5 or above</td>
</tr>
<tr>
<td>English 100</td>
<td>76-82</td>
<td>20-21</td>
<td>540-556</td>
<td>205-219</td>
<td>4.0 or above</td>
</tr>
</tbody>
</table>

As mentioned in the introduction, the ELI environment is relatively unique as the AUC is an English-medium university in an Arabic speaking country, and the ELI is an intensive program designed specifically to raise proficiency levels of academic English. The English 99 students meet five hours a day, four days a week, and their classes are divided into writing, grammar, study skills, reading, and vocabulary. The English 100 students meet three hours a day, four days a week, and their classes are divided into writing, reading, and vocabulary. Due to this specific environment, the results of this study cannot be generalized to all university or academic English contexts. It is the aim of this study however, to encourage further research on this topic and to have the results of this study verified using larger and more diversified samples.
3.2 Sampling Procedure

The participants in this study were selected using convenience sampling. Teachers in the ELI were approached about allowing their class to participate in this study and six graciously agreed. One English 99 class and two English 100 classes were randomly assigned to the experimental group, and the remaining three classes served as the control group. The only criterion for assignment as a treatment or control group was to ensure that there were two English 100 classes in each group and 1 English 99 group in each group in order to avoid having proficiency level as an extraneous variable.

3.3 Materials and Instruments

3.3.1 Materials

A target vocabulary list was first developed for this study. Since the ELI is an EAP program, the vocabulary list included exclusively words from the AWL. Words were selected at random from the AWL, and the Merriam-Webster online thesaurus was used to find any near-synonyms that also appeared on the AWL. The idea was that EFL students often struggle at recognizing the subtle distinctions between near-synonyms, and this can lead to the inaccurate or incorrect usage of words since the majority of synonyms are not fully intersubstitutable. The method of using an online thesaurus to find synonyms is simple and convenient, thus making it likely that this is the way an EFL student would go about finding a synonym.

The original target vocabulary list consisted of 14 word pairs with the expectation that some word pairs would be omitted after piloting. Since a number of vocabulary teachers in the IEP stated that 15-20 words was the average number of words they present to their students in a single lesson, the goal was to keep seven to ten word pairs for the in-
class activities. The final target vocabulary list consists of eight word pairs (16 words) and can be found in Appendix A.

Materials for the in-class vocabulary activity were also designed for this study. The control group received a handout that consisted of the target vocabulary list along with the definition of each word and one or two examples from the Merriam-Webster online dictionary. These definitions and examples were used to complete 16 fill-in-the-blank items that, like on the test, came directly from the COCA corpus (Davies, 1990). The experimental group, on the other hand, received a handout that consisted of the target vocabulary list, the Merriam-Webster definitions of each word, 4-5 concordance lines for each word, and information about phraseology and collocations gathered from the COCA corpus (Davies, 1990). This information was then used to complete the same 16 fill-in-the-blank items as the control group. The key difference between the two activities was that the control group received only the definition and examples from the online dictionary, whereas the experimental group received the definition in addition to concordance lines and information about phraseology and collocations. The handout for the control group is attached in Appendix B, and the handout for the experimental group is attached in Appendix C.

3.3.2 Instruments

A vocabulary test was designed in order to measure the knowledge of the target vocabulary items. The test was modeled after the productive section of Paul Nation’s Vocabulary Levels Test which can be found in *Learning Vocabulary in another Language* (Nation, 2001). However, a slight modification was made in order for the test to better suit the needs of this study. The traditional Vocabulary Levels Test provides
students with the first 2-3 letters of the vocabulary item in order to aid in the recall of the item. Since this study is testing, in addition to overall vocabulary knowledge of the items, distinction between near-synonyms, providing the first few letters of the item would automatically give away which of the two near-synonyms is the correct answer. Therefore, a word bank is provided with the target items in addition to several distracters in order to minimize the effects of guessing. By providing the word bank, the potential answers that the students may use in completing the sentences are limited, thus forcing the students to use specifically the target vocabulary items. This kind of test has traditionally been referred to as a “contextualized word completion recall test” (Dolch, 1931).

Some may question whether this test is actually assessing productive or receptive knowledge of vocabulary. As mentioned in the Definition of Terms section in Chapter I, knowledge of vocabulary cannot simply be divided into either productive or receptive knowledge, but rather it is more of a continuum (D’Anna & Zechmeister, 1991; Henriksen, 1996; Zimmerman, 1997). Generally, it is agreed that the ultimate stage of knowledge of a vocabulary item is the ability to use it accurately in speech or writing also referred to as “free” or “unaided” recall (Kelley & Krey, 1934).

For the purposes of this study, however, it was not plausible to test free production as it does not allow for control of what words are used, thus making it impossible to ensure that students would use the target vocabulary items. The test described above can be classified as a “recall aided by recognition” test using Kelley and Krey’s (1934) long-standing breakdown of vocabulary knowledge test methods and underlying mental processes. A “recall aided by recognition” test requires students to
have more than a simple receptive knowledge of a word (e.g. providing an L1 translation or correctly matching the word with its definition) but does not require them to have reached the ultimate step of free production of the vocabulary item. The ability to identify whether or not a word can be accurately used in a given context is generally classified as the penultimate step or in other words the step directly preceding free production, and this is the knowledge that this study aims to assess (D’Anna & Zechmeister, 1991; Henriksen, 1996; Zimmerman, 1997).

The vocabulary test was piloted with an English 99 class of 10 students that were not participating in the study. By doing so, the group being used to pilot the test was as similar to the participants as possible. Before piloting, the test consisted of 28 fill-in-the-blank items that came directly from the COCA corpus (Davies, 1990). The results of the piloting indicated three problematic questions, and so these words, in addition to their counterpart in the word pair, were removed from the test and the target vocabulary list. In addition to this, questions with an item difficulty score of .80 or above were also omitted which led to the removal of three more word pairs.

It is important to note that there are two “weighting sections” in the test specifications table below. The test was scored once to obtain a score reflective of overall knowledge of the vocabulary items, and it was scored a second time to obtain a score reflective of the ability to distinguish between near-synonyms. The weighting for each method is explained in Table 3.1 and the test, after revisions, is attached in Appendix D.
An open-response questionnaire was also designed to investigate the attitudes of the participants, in the experimental group, towards the use of concordance data in the classroom. The questionnaire was designed as open-response in order to not limit participants’ responses. The questionnaire asked the participants to describe what they liked and did not like about using concordance data, what kind of information they felt they had learned from the data, if this information helped them to learn the vocabulary items, and finally if they would want to use concordance data in their classroom again.

### 3.4 Data Collection Procedures

First and foremost, the pretest was administered to all six intact classes at six different times. The researcher of this study administered all the tests personally in order to ensure consistency of directions. The participants were asked to separate from one another, write their ID numbers on the test instead of using their names and were read the directions for the test (see Appendix D). All participants had 20 minutes to complete the test and were given no extra time or “hints” or “clues” to help them with the test. However, in order to avoid discouragement, the participants were told before starting the
test that some of the words would be unfamiliar to them and were encouraged just to try their best.

Next, the participants were divided into a control group and an experimental group. As mentioned above, one English 99 class and two English 100 classes were assigned to the experimental group, and the remaining classes were assigned to the control group. The week following the pretest, the researcher of this study went to the six intact classes separately in order to explain and monitor the in-class activities. The control group received the activity attached in Appendix B and the experimental group received the activity attached in Appendix C. All groups were read the directions for the activity, given 45 minutes to complete it, and given permission to work in pairs. They were also encouraged to ask questions about anything they did not understand. These questions usually dealt with the meaning of words (other than the target vocabulary words) in the concordance lines. After 45 minutes, the participants were asked to stop working, and the last 15 minutes of class were spent having the participants take turns reading the answers to the 16 questions. It is important to note that this study did not allow the experimental group any more time to complete their activity than the control group. This avoids the possibility that increased attention to the target word items is what caused the results rather than the treatment.

The day immediately following the in-class activity, all the participants took the posttest. Once again, the researcher of this study administered the test to all six classes in order to ensure consistency of administration. The participants were once again asked to separate from one another, write their ID numbers on the test, were read the directions from the test (see Appendix D) and given 20 minutes to complete the test. The posttest
was identical to the pretest previously taken by the participants. While it is possible that familiarity with the test may have influenced results, it would have affected all of the participants’ results equally as they were all equally familiar with the test.

It is also important to note that while all six participating classes did not necessarily take the tests or complete the activity on the same day, there was a minimum of four days between administration of the pretest and the in-class activity for all classes and the posttest was administered the day immediately following the in-class activity for all classes. By keeping the time between administrations as consistent as possible for all participants, the potential effects of memory loss are reduced. Lastly, on the same day as the posttest, the experimental group was asked to complete the questionnaire attached in Appendix E.

3.5 Data Analysis

Parametric inferential statistics was used to analyze the numerical data in this study since the data was in the form of test scores and there was sample size of 50 participants. The analysis simultaneously tested for differences between groups using the mean scores obtained from the “overall vocabulary knowledge” scoring rubric and the mean scores obtained from the “distinction between near-synonyms” scoring rubric. Details on the two scoring rubrics are presented in the Instruments section in Chapter III.

The focus of the analysis was to find any significant differences between the control and experimental group. The control and experimental groups each consisted of one English 99 class and two English 100 classes and were very similar in terms of age, native language, and proficiency level. The control group consisted of 26 participants, ages 17-21, who were all native speakers of Arabic. Nine of the participants were in
English 99 with the TOEFL placement scores discussed in the Participants section in Chapter III and 17 were in English 100 with the TOEFL scores also discussed in the Participants section. The experimental group consisted of 24 participants also ages 17-21 and native Arabic speakers. Nine of the participants were English 99 students and 15 participants were English 100 students.

In this study, the vocabulary activity served as the independent variable (IV) and overall vocabulary knowledge and ability to distinguish between synonyms served as the dependent variables (DV). Therefore, since there was one independent variable, two dependent variables, and two administrations of the same measurement, a multivariate repeated measures ANOVA was the best measurement of change. It was preferable to use a multivariate repeated measures ANOVA, as opposed to multiple univariate repeated measures ANOVAs, in order to control for the probability of making a Type I error.

In summary, the inferential statistical analysis of the numerical data in this study was able to tell us whether or not there was a significant difference in performance on the pretest and posttest between the experimental and control group for either the “total vocabulary knowledge” score or the “distinction between synonyms” score. A significant difference between the control and experimental group would suggest that one of the groups made higher gains in vocabulary knowledge or synonym distinction than the other. A discussion of how these results answer the first and second research questions is included in Chapter V of this study.

Since the questionnaire was open-ended and did not contain a Likert scale, statistics was not used to analyze the data. A table for each question of the questionnaire is presented below in Chapter IV that includes all responses given to the question and the
frequency of each response. The data presented in these tables is discussed descriptively in Chapter IV. This analysis aided in answering the third research question dealing with students’ attitudes toward classroom use of concordance data.
Chapter IV

Results

This study used a pretest/posttest design in order to investigate the effects of using concordance data as a teaching tool on vocabulary acquisition in the ELI at the AUC. Specifically, it investigated the effects of corpus use on overall vocabulary knowledge and ability to distinguish between symptoms. It also used a questionnaire to explore students’ attitudes towards the use of concordance data in their classroom. The results are reported in the following section.

4.1 Results of Vocabulary Test Data

A repeated-measures ANOVA was conducted in order to investigate whether there are any significant differences in vocabulary acquisition between the control group who received a traditional vocabulary lesson and the experimental group who received a corpus-based vocabulary lesson. Table 4.1 presents the mean, standard deviation, and number of participants for the pretest and posttest for each group and for both of the dependent variables: overall vocabulary knowledge and distinction between synonyms. The pretest means for overall vocabulary knowledge are relatively similar for the control and experimental group (M= 8.38, 8.87, SD= 3.359, 3.709, respectively) and so are the pretest means for distinction between synonyms (M= 6.5, 6.7, SD= 4.246, 4.665, respectively). This suggests that the control and treatment groups were relatively similar in ability at the beginning of the experiment.
Table 4.1

Descriptive Statistics

<table>
<thead>
<tr>
<th>Test</th>
<th>Group</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pretest</td>
<td>Control</td>
<td>8.38</td>
<td>3.359</td>
<td>26</td>
</tr>
<tr>
<td>Vocab</td>
<td>Experimental</td>
<td>8.87</td>
<td>3.709</td>
<td>24</td>
</tr>
<tr>
<td>Score</td>
<td>Total</td>
<td>8.61</td>
<td>3.499</td>
<td>50</td>
</tr>
<tr>
<td>Posttest</td>
<td>Control</td>
<td>9.88</td>
<td>3.290</td>
<td>26</td>
</tr>
<tr>
<td>Vocab</td>
<td>Experimental</td>
<td>10.83</td>
<td>3.312</td>
<td>24</td>
</tr>
<tr>
<td>Score</td>
<td>Total</td>
<td>10.33</td>
<td>3.300</td>
<td>50</td>
</tr>
<tr>
<td>Pretest</td>
<td>Control</td>
<td>6.50</td>
<td>4.226</td>
<td>26</td>
</tr>
<tr>
<td>Distinction</td>
<td>Experimental</td>
<td>6.70</td>
<td>4.665</td>
<td>24</td>
</tr>
<tr>
<td>Score</td>
<td>Total</td>
<td>6.59</td>
<td>4.392</td>
<td>50</td>
</tr>
<tr>
<td>Posttest</td>
<td>Control</td>
<td>7.88</td>
<td>4.616</td>
<td>26</td>
</tr>
<tr>
<td>Distinction</td>
<td>Experimental</td>
<td>9.17</td>
<td>4.609</td>
<td>24</td>
</tr>
<tr>
<td>Score</td>
<td>Total</td>
<td>8.49</td>
<td>4.610</td>
<td>50</td>
</tr>
</tbody>
</table>

Table 4.1 also shows an increase in means for the control and experimental group on the posttest for overall vocabulary knowledge (M= 9.88, 10.83, SD= 3.29, 3.312, respectively) and on the posttest for distinction between synonyms (M= 7.88, 9.17, SD=4.616, 4.609, respectively). While it seems, based on Table 4.1, that the increase in mean for both groups from pretest to posttest, the repeated-measures ANOVA must be looked at in order to find out whether this increase is statistically significant or whether one group improved more than another.

Table 4.2 shows the results of the Wilks’ Lambda multivariate test conducted. The results of the multivariate test are included in this section, in addition to the results of the repeated-measures ANOVA, in order to provide a general overview of the data. The multivariate test shows that the overall vocabulary test scores differ significantly within each group, F(1, 47) = 86.565, P<.001, and that the distinction between synonyms scores
also differs significantly within each group, \( F(1,47) = 17.319, P<.001 \). This indicates that the increase in scores from pretest to posttest for both groups is statistically significant.

Table 4.2

**Multivariate test**

<table>
<thead>
<tr>
<th>Effect</th>
<th>Value</th>
<th>F</th>
<th>Hypothesis df</th>
<th>Error df</th>
<th>Sig.</th>
<th>Partial Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Vocabtest</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wilks' Lambda</td>
<td>.352</td>
<td>86.565*</td>
<td>1.000</td>
<td>47.000</td>
<td>.000</td>
<td>.648</td>
</tr>
<tr>
<td><strong>Vocabtest * Group</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wilks' Lambda</td>
<td>1.000</td>
<td>.005*</td>
<td>1.000</td>
<td>47.000</td>
<td>.944</td>
<td>.000</td>
</tr>
<tr>
<td><strong>Distinctiontest</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wilks' Lambda</td>
<td>.731</td>
<td>17.319*</td>
<td>1.000</td>
<td>47.000</td>
<td>.000</td>
<td>.269</td>
</tr>
<tr>
<td><strong>Distinctiontest * Group</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wilks' Lambda</td>
<td>.984</td>
<td>.777*</td>
<td>1.000</td>
<td>47.000</td>
<td>.383</td>
<td>.016</td>
</tr>
<tr>
<td><strong>Vocabtest * Distinctiontest</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wilks' Lambda</td>
<td>.983</td>
<td>.814*</td>
<td>1.000</td>
<td>47.000</td>
<td>.372</td>
<td>.017</td>
</tr>
<tr>
<td><strong>Vocabtest * Distinctiontest * Group</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wilks' Lambda</td>
<td>.959</td>
<td>2.000*</td>
<td>1.000</td>
<td>47.000</td>
<td>.164</td>
<td>.041</td>
</tr>
</tbody>
</table>

*Note. Statistically significant results are presented in bold.*

Table 4.3, showing a repeated-measures ANOVA within subjects, also presents similar results. However, unlike the multivariate test, a repeated-measures ANOVA assumes sphericity meaning it requires equal covariances, in addition to variances, for each level of the within subjects variable (Howell, 2002). Since in this study there are only two groups, and at least three groups are needed to check for sphericity, the results that use the Greenhouse-Geisser correction are used. Once again the results indicate that there is a statistically significant difference from pretest to posttest within each group on the overall vocabulary knowledge, \( F(1,47) = 86.565, P<.001 \), and the distinction between synonyms, \( F(1,47) = 17.319, P<.001 \). However, there is no statistically significant
difference between groups from pretest to posttest either for the overall vocabulary knowledge, F(1,47) = .005, P=.94, or for the distinction between synonyms, F(1,47) =.777, P=.38.

Table 4.3

Tests of Within-Subjects Effects

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
<th>Partial Eta Square</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Vocab test</strong></td>
<td>Sphericity Assumed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>181.398</td>
<td>1</td>
<td>181.398</td>
<td>86.565</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>Greenhouse-Geisser</td>
<td>181.398</td>
<td>1.000</td>
<td>181.398</td>
<td>86.565</td>
<td>.000</td>
</tr>
<tr>
<td><strong>Vocab test * Group</strong></td>
<td>Sphericity Assumed</td>
<td>.010</td>
<td>1</td>
<td>.010</td>
<td>.005</td>
<td>.944</td>
</tr>
<tr>
<td></td>
<td>Greenhouse-Geisser</td>
<td>.010</td>
<td>1.000</td>
<td>.010</td>
<td>.005</td>
<td>.944</td>
</tr>
<tr>
<td><strong>Error(vocabtest)</strong></td>
<td>Sphericity Assumed</td>
<td>98.490</td>
<td>47</td>
<td>2.096</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Greenhouse-Geisser</td>
<td>98.490</td>
<td>47.000</td>
<td>2.096</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Distinction test</strong></td>
<td>Sphericity Assumed</td>
<td>163.454</td>
<td>1</td>
<td>163.454</td>
<td>17.319</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>Greenhouse-Geisser</td>
<td>163.454</td>
<td>1.000</td>
<td>163.454</td>
<td>17.319</td>
<td>.000</td>
</tr>
<tr>
<td><strong>Distinction test * Group</strong></td>
<td>Sphericity Assumed</td>
<td>7.332</td>
<td>1</td>
<td>7.332</td>
<td>.777</td>
<td>.383</td>
</tr>
<tr>
<td></td>
<td>Greenhouse-Geisser</td>
<td>7.332</td>
<td>1.000</td>
<td>7.332</td>
<td>.777</td>
<td>.383</td>
</tr>
<tr>
<td><strong>Error(Distinction test)</strong></td>
<td>Sphericity Assumed</td>
<td>443.577</td>
<td>47</td>
<td>9.438</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Greenhouse-Geisser</td>
<td>443.577</td>
<td>47.000</td>
<td>9.438</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Vocab test * Distinction test</strong></td>
<td>Sphericity Assumed</td>
<td>.504</td>
<td>1</td>
<td>.504</td>
<td>.814</td>
<td>.372</td>
</tr>
<tr>
<td></td>
<td>Greenhouse-Geisser</td>
<td>.504</td>
<td>1.000</td>
<td>.504</td>
<td>.814</td>
<td>.372</td>
</tr>
<tr>
<td><strong>Vocab test * Distinction test * Group</strong></td>
<td>Sphericity Assumed</td>
<td>1.238</td>
<td>1</td>
<td>1.238</td>
<td>2.000</td>
<td>.164</td>
</tr>
<tr>
<td></td>
<td>Greenhouse-Geisser</td>
<td>1.238</td>
<td>1.000</td>
<td>1.238</td>
<td>2.000</td>
<td>.164</td>
</tr>
<tr>
<td><strong>Error(Vocab test*Distinction test)</strong></td>
<td>Sphericity Assumed</td>
<td>29.098</td>
<td>47</td>
<td>.619</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Greenhouse-Geisser</td>
<td>29.098</td>
<td>47.000</td>
<td>.619</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Statistically significant results are presented in bold.
Lastly, Table 5.4 displays the results of a repeated-measures ANOVA of between-subjects effects. The results show that there is no statistically significant difference between groups, \( F(1,47) = .496, P=.485 \). Based on the results of Table 5.3 and 5.4, it can be concluded that both groups performed statistically similarly on both the pretest and posttest for both scores. In other words, neither group made higher gains than the other in either overall vocabulary knowledge or ability to distinguish between synonyms.

Table 4.4

*Tests of Between-Subjects Effects*

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
<th>Partial Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>14198.963</td>
<td>1</td>
<td>14198.963</td>
<td>272.290</td>
<td>.000</td>
<td>.853</td>
</tr>
<tr>
<td>Group</td>
<td>25.861</td>
<td>1</td>
<td>25.861</td>
<td>.496</td>
<td>.485</td>
<td>.010</td>
</tr>
<tr>
<td>Error</td>
<td>2450.884</td>
<td>47</td>
<td>52.146</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

To summarize, the results of the repeated-measures ANOVA indicated that there were statistically significant within-subjects effects. Both the control and the experimental groups made significant gains in terms of overall vocabulary knowledge and ability to distinguish between synonyms. The test of between-subjects effects, however, indicated that there were no significant differences between the two groups which means neither group made higher gains than the other.

4.2 Results of Questionnaire Data

As mentioned in the Data Analysis section in Chapter III, statistics was not used to analyze the verbal data gathered by the questionnaire. The researcher chose to report the percentage of each response for all questions. *Tables 4.5, 4.6, 4.7, 4.8 and 4.9* list the responses provided by the participants along with the frequency and percentage of participants that provided that particular response. Please note that since the
questionnaire was open-ended, participants had the option to give multiple answers for each question, and so the total frequency does not necessarily equal 19 for each question nor does the percentage necessarily equal 100.

Table 4.5 shows us the participants’ responses to the first question on the questionnaire which asked what the participants liked about using concordance lines. Fifty-eight percent of the participants responded that they liked using the concordance lines because they felt that they helped them to understand the meaning more precisely. Twenty-one percent responded that they liked that they were given plenty of examples and that the concordance lines were “useful” and “beneficial”. Participants also liked that they were given information about the collocations and phraseology of each word and that the concordance lines helped them to understand the differences between the synonyms. One participant (5 percent) stated that they did not like anything about the concordance lines and in fact, did not use them. Question 2 provides us with more answers as to what the participants did not like about the concordance lines.

Table 4.5

<table>
<thead>
<tr>
<th>Question 1: Answers</th>
<th>Question 1: Frequency (N=19)</th>
<th>Question 1: Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Helped them to understand exact meaning better</td>
<td>11</td>
<td>58%</td>
</tr>
<tr>
<td>Plenty of examples</td>
<td>4</td>
<td>21%</td>
</tr>
<tr>
<td>Effective/useful/beneficial</td>
<td>4</td>
<td>21%</td>
</tr>
<tr>
<td>Gave us hints (i.e. information about collocations and phraseology)</td>
<td>3</td>
<td>16%</td>
</tr>
<tr>
<td>Helped them to understand differences between synonyms/allowed for comparisons</td>
<td>2</td>
<td>11%</td>
</tr>
<tr>
<td>Nothing/did not use</td>
<td>1</td>
<td>5%</td>
</tr>
</tbody>
</table>

Table 4.6 presents the participants’ responses to Question 2 which dealt with things they disliked about using the concordance lines. Seventy-four percent of the
participants responded with “nothing” while 11 percent wanted more exercises to accompany the concordance lines, 5 percent felt they took too long to read, and 5 percent did not use them. The responses to this question would suggest that the majority of the students enjoyed working with the concordance lines.

Table 4.6

*Question 2: What did you dislike about using the concordance lines to complete the in-class activity?*

<table>
<thead>
<tr>
<th>Question 2: Answers</th>
<th>Question 2: Frequency (N=19)</th>
<th>Question 2: Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nothing</td>
<td>14</td>
<td>74%</td>
</tr>
<tr>
<td>Not enough exercises</td>
<td>2</td>
<td>11%</td>
</tr>
<tr>
<td>Took too much time to read</td>
<td>1</td>
<td>5%</td>
</tr>
<tr>
<td>Did not use them</td>
<td>1</td>
<td>5%</td>
</tr>
</tbody>
</table>

The responses to Question 3 are presented in *Table 4.7* below and deal with what kind of information the participants felt they learned about the words from the concordance lines. It is important to find out not just what the participants liked and disliked about using the concordance lines but also what they felt they learned. Fifty-three percent of participants reported that they had learned how to use the words in a sentence and 47 percent responded that they learned the meaning of the words. Twenty-two percent also felt that they had learned the differences between the synonyms while 16 percent responded that they had learned information about the words’ collocations and phraseology.
Table 4.7

*Question 3: What kind of information did you learn about each word based on the concordance lines?*

<table>
<thead>
<tr>
<th>Question 3: Answers</th>
<th>Question 3: Frequency (N=19)</th>
<th>Question 3: Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word uses/how to use the word in a sentence</td>
<td>10</td>
<td>53%</td>
</tr>
<tr>
<td>The meaning of the word</td>
<td>9</td>
<td>47%</td>
</tr>
<tr>
<td>Difference between the synonyms</td>
<td>4</td>
<td>21%</td>
</tr>
<tr>
<td>Hints (i.e. information about collocations and phraseology)</td>
<td>3</td>
<td>16%</td>
</tr>
</tbody>
</table>

*Table 4.8* shows that 95 percent of the participants responded “yes” when asked if they felt the concordance lines helped them to learn the vocabulary items. Specifically, they stated that the concordance lines helped them to learn the differences between the synonyms (21%), how to use the words (53%) and the definition/meaning of the words (47%). One participant (5%) responded that the concordance lines did not help them to learn the vocabulary items because they were already familiar with the words.

Table 4.8

*Question 4: Do you think the concordance lines and in-class activity helped you to learn the vocabulary words? If yes, in what ways? If no, why not?*

<table>
<thead>
<tr>
<th>Question 4: Answers</th>
<th>Question 4: Frequency (N=19)</th>
<th>Question 4: Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>18</td>
<td>95%</td>
</tr>
<tr>
<td>Learned differences between the synonyms</td>
<td>10</td>
<td>53%</td>
</tr>
<tr>
<td>Learned how to use words</td>
<td>6</td>
<td>32%</td>
</tr>
<tr>
<td>Learned definition/meaning of the words</td>
<td>5</td>
<td>26%</td>
</tr>
<tr>
<td>Gave many examples</td>
<td>4</td>
<td>21%</td>
</tr>
<tr>
<td>No</td>
<td>1</td>
<td>5%</td>
</tr>
<tr>
<td>Already knew the words</td>
<td>1</td>
<td>5%</td>
</tr>
</tbody>
</table>

Lastly, *Table 4.9* shows that 95 percent of the participants would like to use concordance lines in class again. When asked to explain their answer, they responded that they wanted to use the concordance lines again because they had helped them to understand the words better (42%), improved their language (16%), helped with learning
the differences between the synonyms (11%) and was an easier way to learn vocabulary (11%). One participant left this question blank.

Table 4.9

*Question 5: Would you want your teacher to use concordance lines in class again? Why or why not?*

<table>
<thead>
<tr>
<th>Question 5: Answers</th>
<th>Question 5: Frequency (N=19)</th>
<th>Question 5: Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>18</td>
<td>95%</td>
</tr>
<tr>
<td>Helped to understand the words better</td>
<td>8</td>
<td>42%</td>
</tr>
<tr>
<td>Improved their language</td>
<td>3</td>
<td>16%</td>
</tr>
<tr>
<td>Helped with distinction between synonyms</td>
<td>2</td>
<td>11%</td>
</tr>
<tr>
<td>Made learning vocabulary easier</td>
<td>2</td>
<td>11%</td>
</tr>
<tr>
<td>Blank</td>
<td>1</td>
<td>5%</td>
</tr>
</tbody>
</table>

To summarize, the results of the questionnaire indicated that the participants of this study held a positive attitude towards the use of the concordance lines. Participants reported that they liked using the concordance lines because they helped them to understand the meaning of the target word items, they provided many examples and they helped them to learn the differences between the synonyms. Impressively, 74% of the participants reported that there was nothing that they disliked about using the concordance lines and 95% indicated that they felt the lines had helped them to learn the target word items and that they would want to use concordance lines in class again.
Chapter V

Discussion and Conclusion

5.1 Discussion

This study aimed to investigate the effects of classroom corpus use on vocabulary acquisition. To look into these issues, a vocabulary test measuring overall lexical knowledge and ability to distinguish between synonyms was used. In addition, a questionnaire was also administered in order to explore students’ attitudes towards the use of concordance lines in their classroom.

Repeated-measures ANOVA was used to analyze data obtain from the vocabulary test. The results of the repeated-measures ANOVA showed that while the participants within each group improved from pretest to posttest, there was no significant difference across the control and experimental groups. In other words, the experimental group did not make higher gains than the control group which suggests that both the control lesson and the experimental lesson were equally beneficial to the participants in terms of vocabulary acquisition as defined in the Definition of Terms section in Chapter I.

5.2 First Research Question

To answer the first research question, implementing the use of concordance data in the classroom for a one-hour lesson did have an effect on vocabulary acquisition. The experimental group made significant gains in vocabulary knowledge after the use of the concordance data in their classes. The effects, however, were not statistically different than the effects of using the traditional ELI method of teaching vocabulary. These results seem to both confirm and contradict the findings of Cobb (1997). Cobb found that participants who used a concordancer scored significantly higher from pretest to posttest.
which would suggest, like the results of this study, that use of concordance lines does have an effect on vocabulary acquisition. However, Cobb’s results indicated that the participants who used the concordancer scored significantly higher than those who did not on the same vocabulary test. Possible reasons for why this study did not find a significant difference between the control and treatment group are discussed in the following section.

5.3 Second Research Question

To answer the second research question, using concordance data in the classroom also had an effect on the ability of participants to distinguish between synonyms. Once again, the experimental group made significant gains in the ability to distinguish between synonyms, but these gains were not significantly different from the gains made by the control group. These results seem to confirm findings of Kaur and Hegelheimer (2005), whose results showed that participants using a concordancer in addition to an online dictionary did not score significantly higher on vocabulary tasks than participants using only the online dictionary. However, they did find that the participants who used the concordancer had a higher transfer rate of the target words to their writing. Use of target words in the participants’ writing was not investigated in this study.

5.4 Third Research Question

Finally, the third research question deals with students’ attitudes towards classroom use of the concordance lines. Previous studies such as Varley (2009) found that students, while finding the concordancing software useful, disliked using it themselves because it was “confusing” and “overwhelming”. In addition, Kaur and Hegelheimer (2005) found that students were hesitant to make use of a concordancer
when given an option. For this reason, this study did not have the participants use the concordancing software themselves, but rather they were presented with pre-selected concordance lines. This seems to have led to a more positive attitude towards the use of concordance data than was reported in Varley (2009).

The participants in this study overwhelmingly (95%) reported that they felt the concordance lines helped them in learning the vocabulary items and that would want to use them again in class. Seventy-four percent of the participants reported that they disliked nothing about using the concordance lines and the other participants responded that they wanted more exercises to accompany the concordance lines or they wanted more time to read them. It seems that by limiting the concordance data to 4-5 pre-selected lines, the participants did not feel overwhelmed or confused by the data. Since the results of this study suggest that using concordance lines is just as effective a method of teaching vocabulary as what is currently being used in the ELI, teachers can use concordance data in class as a way of varying activities to make class more exciting and less monotonous for students.

5.5 Implications of Findings

It has been suggested in numerous studies that corpora have the potential for major implications in the TESOL world. Wang and Good (2007) have suggested that textbooks be written in accordance to the findings of corpora in order to reflect authentic language use. Shirato and Stapleton (2007) have suggested that curricula be redesigned in order to incorporate corpora in the classroom. Biber et al. (1998) have suggested that corpora be used in order to analyze discrepancies between L1 and L2 target language use and consequently the areas of discrepancy should be focused on in the classroom.
The implications of the findings of this study are primarily on classroom teaching and material use. The results of this study indicate that providing students with contextual examples of the target words has a significant effect on vocabulary acquisition, and students themselves see the value in studying the words in context and in fact, enjoy it. This study, in addition to Collins (2010), Mason et al. (2009), and Verhallen and Bus (2010), suggests that teachers provide their students with contextual examples of vocabulary words when teaching these words to their students.

In addition to this, the participants in this study repeatedly stated that they enjoyed having multiple examples of each word because it helped them to truly understand the meaning of the word and to understand the differences between words. Therefore, in addition to Zahar et al. (2001), this study suggests that teachers provide students with multiple examples of each word in context.

In order to truly change the way vocabulary is taught, changes are needed at the curriculum-design level. The objective of a vocabulary class should no longer be to memorize the target vocabulary items and the L1 translation, but rather to be able to fully recognize and produce the items. In order to achieve this, it may be advisable for vocabulary to be integrated more with other language skills such as reading and writing instead of being taught as a separate language skill.

In terms of pedagogy, the results of this study support the idea that the classroom should be a more learner-centered environment. The use of corpus lines in the classroom allows students to explore the language for themselves and to discover meanings of words, collocations, and phraseology (Johns, 1986; Breyer 2009). Meanwhile, it is easy to couple this exploration of language with an explicit vocabulary task, thus reinforcing
the learning of vocabulary items and ensuring students to not infer the incorrect meaning of words (Walters & Bozkurt, 2009).

5.6 Limitations of the Study

It is, of course, possible that a number of extraneous variables influenced the results of this study. First and foremost, the sample size of this study was quite small. It is possible that with a significantly larger sample size, the results may have differed. Further studies should be conducted with larger sample sizes to investigate the effects of classroom use of concordance data on vocabulary acquisition in addition to other language skills.

Another factor that may have influenced the results is motivation. The participants in this study knew that their scores on the test and participation in the lesson would in no way affect them and this appeared to have a demotivating affect on their performance. Evidence supporting this idea is that two participants, one in the control group and one in the treatment group, performed significantly worse on the posttest. It appeared they had written just any answer in order to complete the test and be allowed to leave class. Also, one participant wrote in the questionnaire that they did not even use the concordance lines. It appeared at times that some just wanted to complete the activity as quickly as possible and did not put any great effort into getting as much out of the lesson as they could. Perhaps if concordance data were used in classrooms by the students’ actual teachers and as a “real” part of their class, they would be more motivated to invest their time and effort into using the concordance lines. If this were the case, the results may differ as motivation has long been known to have a powerful effect on language learning (Gardner & Smythe, 1975; Oxford & Shearin, 1994).
Another potential limitation was duration of treatment. It was not possible for the purposes of this study to implement a semester-long treatment of corpus based lessons. The participants of this study were only exposed to concordance lines on one occasion for one hour, and so it is possible that the use of concordance data in the classroom may have effects on vocabulary acquisition that were not seen in this study.

Lastly, just because the results did not show a statistically significant difference between the performance of the control and experimental group on a “recall aided by recognition” test, does not mean that the experimental group did not make higher gains in another aspect of vocabulary acquisition. Free production of vocabulary items was not tested in this study and Kaur and Hegelheimer (2005) found that students who used concordancing software had a higher and more accurate transfer rate of target vocabulary items to their own writing. Also, this study did not look into long-term retention of vocabulary items and it is possible that the long-term effects of using concordance data may have been different than the effects of the control lesson.

There is still a great need for further research to be conducted in the area of corpus linguistics specifically on how we can implement corpus use in our classrooms. Future studies should address the limitations encountered in this study, mainly the small sample size, the lack of motivation of the participants, and the short duration of the treatment. They should also investigate potential effects of classroom corpus use on free production of vocabulary items, long-term retention of vocabulary, and other language skills such as grammar as some studies have reported promising results of corpus-based approached on grammar (Liu & Jiang, 2010). It would also be advisable for future
studies to include an interview as part of their methodology in order to get participants to expand and explain their responses on the questionnaire.

5.7 Conclusion

The results of this study suggest that using concordance data in the classroom is an equally effective activity for teaching vocabulary as a more traditional “fill-in-the-blank” activity. Teachers should use a variety of activities with their students to keep the classroom an exciting place and the results of this study suggest that the participants enjoyed using the concordance lines. Specifically, they liked being given multiple contextual examples in addition to getting information about each word’s collocations and phraseology. Eighteen out of 19 participants stated that the concordance lines helped them to learn the vocabulary items and that they would like their teachers to use concordance lines in the classroom again in the future. Concordance lines are one more tool that teachers have available to them to help with the teaching of vocabulary and hopefully future studies will reveal even further benefits of using corpus data in the classroom.
References


Collins, M. (2010). ELL preschoolers’ English vocabulary acquisition from storybook


Hulstijn, J. (1992). Retention of inferred and given word meanings: Experiments in


## Appendix A

### List of Target Word Pairs and Parts of Speech

1. A.) context (N)  
   B.) environment (N)
2. A.) occur (V)  
   B.) happen (V)
3. A.) acquire (V)  
   B.) obtain (V)
4. A.) generate (V)  
   B.) create (V)
5. A.) fluctuation (N)  
   B.) change (N)
6. A.) indicate (V)  
   B.) signify (V)
7. A.) primarily (Adv)  
   B.) predominantly (Adv)
8. A.) guarantee (V)  
   B.) assure (V)
Appendix B
Control Group Handout

Synonyms and the Academic Word List

Student ID: ____________________

Directions: Please read the following definitions and examples carefully and use them to fill-in-the-blank with the correct synonym.

“Signify” versus “Indicate”

Signify (V)

Definition: to be a sign of

Examples:
1.) He gave her a diamond ring to signify his love
2.) A check mark next to your name signifies that you have met all the requirements.

Indicate (V)

Definition: to point out or point to; to be a sign, symptom or index of

Examples:
1.) Our records indicate a depth of 3,000 feet here.
2.) The map indicates where the treasure is buried.

Exercises:

1.) I had the students move their hands up and down to _____________ high and low.
2.) Data _____________ that, for most people, 1 night with 2 hours less sleep affects…
“Context” versus “Environment”

**Context (N)**
Definition: the parts of discourse that surround a word and throw light on its meaning; the interrelated conditions in which something exists or occurs

Examples:
1.) We need to look at the event within the larger context of world history.
2.) We need to consider these events in context.

**Environment (N)**
Definition: the circumstances, objects, or conditions by which one is surrounded

Examples:
1.) He grew up in a loving environment.
2.) Many plans are unable to survive in such a harsh environment.

**Exercises:**
3.) Turning to a larger ____________, I wonder if…
4.) We wanted…them to have a safer work ____________.

“Occur” versus “Happen”

**Occur (V)**
Definition: to come into existence; to come to mind

Examples:
1.) The disease tends to occur in children under the age of five.
2.) No one was ready for what was about to occur.

**Happen (V)**
Definition: to occur by chance; come into being
Examples:
1.) Something like that was bound to happen sooner or later.
2.) You never know what’s going to happen when they get together.

Exercises:
5.) Hilary Clinton…has the experience to actually make change ____________.
6.) There are no rules or plot. A series of random events ____________.

“Acquire” versus “Obtain”

Acquire (V)
Definition: to get as one’s own

Examples:
1.) The two ships were acquired by the navy after the war.
2.) The team acquired three new players this year.

Obtain (V)
Definition: to gain or attain usually by planning action or effort

Examples:
1.) The information may be difficult to obtain.
2.) We obtained a copy of the original letter.

Exercises:
7.) To ____________ a certificate for this permit, I spent less than an hour with my classmates in the shooting range.
8.) Students…__________ process skills, such as higher level thinking skills.

“Create” versus “Generate”

Create (V)
Definition: to bring into existence; to produce through imaginative skills

Examples:
1.) Several new government programs were created while she was governor.
2.) The president has announced a plan to create new jobs.

Generate (V)
Definition: to bring into existence

Examples:
1.) Windmills are used to generate electricity.
2.) We hope to generate some new ideas at the meeting.

Exercises:
9.) How do poor people ____________ wealth?
10.) She continues to ____________ paintings and exhibit them in galleries around the world.

“Fluctuation” versus “Change”

Fluctuation (N)
Definition: to shift back and forth uncertainly (Definition for verb form: Fluctuate)

Examples:
1.) His popularity has fluctuated during his term in office.
2.) In the desert, the temperature fluctuates dramatically.

Change (N)
Definition: the act, process or result of changing

Examples:
1.) There has been little if any change in her daily routine.
2.) We’ve had to make a slight change in the schedule.

Exercises:

11.) Much _______________ in weight may be just due to the amount of water the body is carrying.

12.) Destroying the wicked never seemed to _______________ the world.

“Assure” versus “Guarantee”

Assure (V)
Definition: to inform positively; to make certain the coming or attainment of

Examples:
1.) I can assure you that you won’t be disappointed.

Guarantee (V)
Definition: to assert confidently

Examples:
1.) The washer is guaranteed against defects for one year.
2.) They guarantee that the diamonds they sell are top quality.

Exercises:

13.) The United States is determined to _______________ our allies that we are going to be reliable in helping them.

14.) If they could _______________ that no one would kill me, I would go back tomorrow.

“Predominantly” versus “Primarily”

Predominantly
Definition: for the most part
Examples:

1.) No example.

**Primarily**

**Definition:** for the most part

Examples:

1.) Ketchup is **primarily** made of tomatoes.

2.) The university was **primarily** an agricultural college when it was founded over two centuries ago.

**Exercises:**

15.) Tal Afar (in Iraq) is a _____________ Shiite city.

16.) The first interview concentrated _____________ on premarital experiences.
Appendix C

Experimental Group Handout

Synonyms and the Academic Word List

Directions: Please read the following definitions and examples carefully and use them to fill-in-the-blank with the correct synonym.

“Signify” versus “Indicate”

Signify (V)

Definition: to be a sign of

Hint: Signify + the/a/noun

Examples:

1.) For many Mexicans, its name has come to signify terror and bloodshed.

2.) …it did not signify the beginning of the end.

3.) He wanted to convince the nation that words like "democracy" and "freedom" had come to signify their opposites.

4.) I use the term "identity" to signify a person's understanding of who s/he is…

Indicate (V)

Definition: to point out or point to; to be a sign, symptom or index of

Hint: “Indicate” is very common in academic writing

Most common words: findings/studies/numbers + indicate + that

Examples:

1.) The findings indicate that the most academic success occurred with the children who received the most intensive intervention

2.) A score of 15 would indicate all questions were answered correctly
3.) …there was little evidence to indicate that it was effective

4.) Abundant research data indicate that there are both physiological and physical responses during music listening

**Exercises:**

5.) I had the students move their hands up and down to ____________ high and low.

6.) Data ____________ that, for most people, 1 night with 2 hours less sleep affects…

**“Context” versus “Environment”**

**Context (N)**

**Definition:** the parts of discourse that surround a word and throw light on its meaning; the interrelated conditions in which something exists or occurs

**Hint:** Most common words: larger/narrow/limited/historical/cultural + context

Most commonly used in the form “In/Within the context of” or “in/out of context”

**Examples:**

1.) In the context of food production…

2.) In this context, the most important differences have to do with…

3.) Turkey’s relations with Syria should be understood within the general context of Turkish-Arab relations…

4.) This is not to excuse her behavior, but only to try to see it in a broader historical context.

5.) …pulling quotes out of context tells only part of the story.
Environment (N)

Definition: the circumstances, objects, or conditions by which one is surrounded

Hint: Most common words: learning/classroom/safe/supportive/hostile + environment

Examples:

1.) If students are hearing impaired, a classroom environment that encourages discussions and the sharing of ideas may not be the best fit for them.

2.) … the creation of a non-threatening supportive environment.

3.) …resources that may have potentially beneficial or harmful effects on the environment.

4.) The only requirements are time and the belief that surrounding middle schoolers with a beautiful environment is worthwhile.

Exercises:

7.) Turning to a larger _______________, I wonder if…

8.) We wanted…them to have a safer work ______________.

“Occur” versus “Happen”

Occur (V)

Definition: to come into existence; to come to mind

Hint: Three times more common in academic writing, more formal than “happen”

- Usually used with some degree of certainty: likely/not likely/frequently/will/will not + occur

- Most common words: events/changes/injuries + occur

Examples:
1.) The majority of global population growth in the coming decades will **occur** in those countries where gender disparities are the greatest.

2.) An estimated 80 percent of minor injuries **occur** at home.

3.) Although attitude change did not **occur**…

4.) It did not **occur** to me that there was a very good reason why I did not hear this term.

**Happen (V)**

**Definition:** to occur by chance; come into being

**Hint:** Most common preposition: “Happen + to” (i.e. something happens to someone)

You can also “let” or “make” something happen.

**Examples:**

1.) Of course, such things are not very likely to **happen**.

2.) I don't want to see that **happen**.

3.) Whatever measures it takes, I'm determined to never let it **happen to** me again.

4.) When people have power, they have the ability to **make things happen**.

5.) Though accidents will always **happen**…

**Exercises:**

5.) Hilary Clinton…has the experience to actually make change _______________.

6.) There are no rules or plot. A series of random events _______________.

“**Acquire**” versus “**Obtain**”

**Acquire (V)**

**Definition:** to get as one's own
**Hint:** Most common words: Acquire + skills/knowledge/weapons/property (usually implies a long process)

**Examples:**

1.) Students are asked to acquire the knowledge, attitudes, and interpersonal skills to help them understand and respect their self and others.

2.) Nathanson formed a plan to acquire ideally located, large theaters with a view to refurbishing them.

3.) Reading comprehension is the most important skill one can acquire in school.

4.) We would like our graduates to acquire a strong sense of right and wrong.

**Obtain (V)**

**Definition:** to gain or attain usually by planning action or effort

**Hint:** Most common form: “In order to obtain/In order for (someone) to obtain”

Most common words: Obtain + information, approval, support (necessary things)

**Examples:**

1.) The instructor can obtain assistance from the school's media specialist.

2.) …human resources companies use bribes to obtain required documents and clearances.

3.) The United States failed to obtain Security Council authorization for the war in Iraq.

4.) The sources used to obtain the data are in appendix 2.

5.) They required anyone who sold alcohol to obtain a license or pay a fine.
Exercises:

7.) To ______________ a certificate for this permit, I spent less than an hour with my classmates in the shooting range.

8.) Students…______________ process skills, such as higher level thinking skills.

   “Create” versus “Generate”

Create (V)

Definition: to bring into existence; to produce through imaginative skills

Hint: Most common words: Create + environment/jobs/opportunities/climate

Examples:

1.) We have to create our own book, just like this author did.

2.) Teachers who care and respect their students create an environment that maximizes learning.

3.) This could create a threat to both the physical and the spiritual health…

4.) The primary objective of the government should be to create jobs for the people.

5.) We also introduced methods that would create opportunities for research projects.

Generate (V)

Definition: to bring into existence

Hint: Most common words: Generate + electricity/ideas/solutions/hypotheses

Examples:

1.) We must generate awareness regarding this disease amongst common people.

2.) In this way, children generate questions individually and as a class.

3.) The purpose is to generate interest in the issue.
4.) Since developing tourism and its related products will also generate increased revenue…

5.) Sweden has been heavily dependent on nuclear power to generate electricity.

Exercises:

9.) How do poor people ______________ wealth?

10.) She continues to ______________ paintings and exhibit them in galleries around the world.

“Fluctuation” versus “Change”

Fluctuation (N)

Definition: to shift back and forth uncertainly (Definition for verb form: Fluctuate)

Hint: Much LESS common than “change”.

Most common words/form: “Fluctuation + in/of + temperatures/prices” (things that normally change frequently)

Examples:

1.) …rapid fluctuation in world food prices.

2.) Countries…have sought new strategies to protect themselves from exchange rate fluctuation.

3.) We have no information on their impact on the fluctuation in birth rates.

4.) This causes temperature fluctuation in caves.

5.) An average body temperature is 37 degrees Celsius, but some fluctuation is normal.

Change (N)

Definition: the act, process or result of changing
**Hint:** Most common words/form: “A change + in/on/of + position/attitude/leadership”

**Examples:**

1.) …have resulted in the **change** in the position and role of women in society.

2.) There was minimal **change** in teachers’ responses between the initial and final survey.

3.) What was the cause of the Syrian **change** of policy?

4.) We are working together to make a **change** on the global level.

5.) A **change** in education is a further necessity.

**Exercises:**

11.) Much _______________ in weight may be just due to the amount of water the body is carrying.

12.) Destroying the wicked never seemed to _______________ the world.

> “Assure” versus “Guarantee”

**Assure (V)**

**Definition:** to inform positively; to make certain the coming or attainment of

**Hint:** Most common form: “Assure + someone + that” or “Assure that”

**Examples:**

1.) By preselecting the sites, teachers can **assure** that students are accessing accurate and relevant information.

2.) You have to be able to **assure** people that new development is not going to affect them negatively.

3.) In order to **assure** a quality education…

4.) The soldiers are doing their job: **assure** the security of the embassy.
**Guarantee (V)**

Definition: to assert confidently

Hint: Most common form: “Guarantee + that” is VERY frequent.

Most common words: “Guarantee + rights, security, success, freedom”

Examples:

1.) Good work does not guarantee success.

2.) The purpose of the program is to guarantee that each kindergarten student hears a different book read each day.

3.) the U.S. and Iraqi armies cannot guarantee security to all of Iraq simultaneously.

4.) They do not guarantee that he will not make a regrettable decision.

**Exercises:**

13.) The United States is determined to ____________ our allies that we are going to be reliable in helping them.

14.) If they could ______________ that no one would kill me, I would go back tomorrow.

“**Predominantly” versus “Primarily”**

**Predominantly**

Definition: for the most part

Hint: Most common words: Predominantly Muslim/Christian/white/male/female + population/neighborhood/community.

Examples:

1.) …Brazil has been a **predominantly** Christian country.
2.) The group's soldiers previously were **predominantly** Arab; today, they are largely Pakistani.

3.) After teaching in a **predominantly** white school, I have found that…

4.) The **predominantly** Muslim population earned a living collecting…

**Primarily**

**Definition:** for the most part

**Hint:** MUCH more common than “predominantly”

Most common words/form: concerned/responsible/interested + primarily +

* in/on/from (prepositions)

1.) …will be located **primarily** in North America

2.) International law arguments against the settlements have rested **primarily** upon two sources.

3.) Many rare species are limited **primarily** by the availability of suitable habitat

4.) I have been interested **primarily** in one thing, and that is discrimination.

**Exercises:**

15.) Tal Afar (in Iraq) is a _______________ Shiite city.

16.) The first interview concentrated _______________ on premarital experiences.
Appendix D

Vocabulary Test

Student ID: ______________________

Level (circle one): English 99/English 100

Directions: Please complete the following sentences with the most appropriate word from the word bank. You cannot use a word more than once. You will have 20 minutes to complete the test.

Note- There are extra words in the word bank so you will not use all of the words.

Word Bank

<table>
<thead>
<tr>
<th>acquire</th>
<th>obtain</th>
<th>happen</th>
<th>convert</th>
</tr>
</thead>
<tbody>
<tr>
<td>environment</td>
<td>generate</td>
<td>context</td>
<td>fluctuation</td>
</tr>
<tr>
<td>signify</td>
<td>indicate</td>
<td>change</td>
<td>assure</td>
</tr>
<tr>
<td>guarantee</td>
<td>source</td>
<td>occur</td>
<td>primarily</td>
</tr>
<tr>
<td>initiate</td>
<td>create</td>
<td>analysis</td>
<td>predominantly</td>
</tr>
</tbody>
</table>

1.) The findings are discussed within the ___________ of helping educators to better meet students’ educational needs.

2.) Over the past decade, educators have experimented with ways to help students to start to ___________ language naturally.

3.) While there is a natural ___________ in sea levels, here the trend has been up for as long as they’ve been able to measure it.

4.) I can ___________ you that we will take all actions that may be necessary to protect the province.

5.) With such a stressful ___________, physical and mental illnesses often afflict talented teachers turning them away from teaching and coaching professions.

6.) What would ___________ if you contacted local businesses in your school’s neighborhood and asked them to come to school to discuss how their business functions? Do you think they would say yes?

7.) Not until his last years did it ___________ to me how much was missing from his life and from our relationship.

8.) He was ready for a/an ___________ in his life.

9.) My writing doesn’t ___________ a big income, but it’s what I love to do most.
10.) It is important to _______________ an atmosphere of cohesiveness in the classroom.

11.) In field experiments, researchers must _______________ permission to ask questions.

12.) Results showed that the sample was _______________ middle and lower-middle class. There were only a few participants from the upper class.

13.) He thrust his chest forward to _______________ bravery.

14.) We _______________ that on every flight there will always be at least 10 seats sold at $10.

15.) The narrow description focuses _______________ on the nature of relationships between two people.

16.) Recent reports _______________ that 75 percent of children diagnosed with various forms of cancer in the United States are expected to survive their disease and treatment.
Appendix E

Questionnaire

Level (circle one): English 99/English 100

Please answer the following questions as honestly as you can and include as much detail as possible.

1.) What did you like about using the concordance lines to complete the in-class activity?

2.) What did you dislike about using the concordance lines to complete the in-class activity?

3.) What kind of information did you learn about each word based on the concordance lines?
4.) Do you think the concordance lines and in-class activity helped you to learn the vocabulary words? If yes, in what ways? If no, why not?

5.) Would you want your teacher to use concordance lines in class again? Why or why not?
Appendix F

Consent Form

Title of Project: Investigating the Effects of Using Concordance Data on Vocabulary Acquisition in an Egyptian English for Academic Purposes Setting

Name of Researcher: Jenna Steiner

Please circle yes or no for each item below:
(Note-you must answer “yes” to all items in order to participate in the study.)

- I confirm that I understand the information given to me about the study. □ □
- I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily. □ □
- I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason. □ □
- I understand that data collected during the study, may be looked at by responsible individuals where it is relevant to my taking part in this research. I give permission for these individuals to have access to my test results. □ □
- I agree to take part in the above research study. □ □

Name of Participant __________________________

Date __________________________

Signature __________________________

Name of Person taking consent (if different from researcher) __________________________

Date __________________________

Signature __________________________

Researcher __________________________

Date __________________________

Signature __________________________