

TEACHING ACADEMIC VOCABULARY WITH CORPORA:
STUDENT PERCEPTIONS OF DATA-DRIVEN LEARNING

Stephanie A. Balunda

Submitted to the faculty of the University Graduate School
in partial fulfillment of the requirements
for the degree
Master of Arts
in the Department of English,
Indiana University

December 2009

Accepted by the Faculty of Indiana University, in partial fulfillment of the requirements for the degree of Master of Arts.

Julie A. Belz, Ph.D., Chair

Master's Thesis
Committee

Ulla M. Connor, Ph.D.

Thomas A. Upton, Ph.D.

ACKNOWLEDGEMENTS

I would like to express my sincerest thanks to my thesis committee members, Professors Julie A. Belz, Ulla M. Connor, and Thomas A. Upton. I wish also to sincerely thank Professor Aye-Nu Duerksen for giving me the opportunity to work with the students in the EAP program, and for her continual support during the appointment. Thanks to Suzan Stamper for helping me with all the logistics of the project, and also to the G011 students whose comments and insights fill many of these pages and without which I would have had nothing to say.

Finally, I am most grateful and thankful for the support and love of my father, Richard, who works hard so I can spend my time thinking about and studying language.

TABLE OF CONTENTS

Chapter One Introduction	1
Chapter Two Literature Review.....	6
2.1 History of corpus linguistics	6
2.2 Corpora in linguistic research and language teaching	8
2.3 Data-driven learning (DDL) and vocabulary instruction.....	13
2.4 Formulaic language.....	16
2.4.1 Reasons for teaching formulaic language.....	18
2.4.2 Difficulties of teaching formulaic language	20
Chapter Three The Study	22
3.1 Methods.....	23
3.1.1 Subjects and context	23
3.1.2 Materials	23
3.1.3 Training sessions.....	26
3.1.4 The experiment	28
3.2 Results.....	31
3.2.1 Results for research question one	31
3.2.2 Results for research question two	32
3.2.3 Results for research question three	37
3.2.4 Results for research question four.....	38

Chapter Four Discussion and Conclusion.....	40
4.1 Discussion.....	40
4.2 Conclusion	54
Appendices.....	55
References.....	65
Curriculum Vitae	

TABLES

Table 1	Corpus of G011 Sociology Texts Word Count.....	24
Table 2	Student Comments on Learning Parts of Speech in the Dictionary vs. DDL.....	34
Table 3	Student Comments on the Connection between Word Forms and Meaning	34
Table 4	Student Comments on Learning the Meaning and Function of Words.....	35
Table 5	Student Comments on DDL vs. Guessing the Meaning of Words from Context.....	36
Table 6	Student Comments on the Drawbacks of DDL	37
Table 7	Student Perceptions of what the DDL Activity was Designed to Teach	38

FIGURES

Figure 1	Sample Question from AWL Vocabulary Pre-Test.....	29
Figure 2	Student Perceptions of the Difficulty Level of the DDL Activity	31

APPENDICES

Appendix A	Student Background Information	55
Appendix B	List of AWL Headwords for G011 Course	56
Appendix C	Introduction to Formulaic Language Hand-Out	57
Appendix D	DDL Training Activity.....	59
Appendix E	Vocabulary and Grammar Questionnaire	60
Appendix F	Experiment Session DDL Activity (Group 1)	61
Appendix G	Experiment Session DDL Activity (Group 2).....	62
Appendix H	Post-Activity Interview Protocol.....	63

CHAPTER ONE

INTRODUCTION

Since the mid-1960's, linguists have been using computerized corpora, or large principled collections of electronic texts, to facilitate descriptive analyses of language. In those early years, using corpora to study language (corpus linguistics) was little more than a technology. However, over the years, corpus linguists, as they have come to be called, have outlined principles to justify how and why this technology can be used by linguists. As a result, corpus linguistics has developed into a methodology within the field of linguistics. One of the primary principles which corpus linguists assert is that the study of language should be primarily an empirical endeavor and descriptions and theories of language should be based on the systematic observation of actual language behavior.

The earliest uses of corpora were largely restricted to research. However, advancements in computer technology have permitted greater access to corpora. This democratization of technology has led to an explosion of research by linguists of every persuasion who have realized the remarkable potential of corpora and accompanying software to facilitate and supplement most linguistic studies.

At the same time many linguists were becoming interested in using corpora to do linguistic research, which led to many new avenues of linguistic enquiry and more complete and accurate descriptions of language structure and use than ever before, Communicative Language Teaching (CLT) arrived as the dominant paradigm for teaching second and foreign languages in the U.S. in the late 1970s and early 1980s. Because this approach emphasized exposing learners to authentic language in context,

gave form-focused instruction a warranted place in the curriculum, and encouraged the use of inductive learning techniques where appropriate, linguists soon began to see the overlap between the methods used by corpus linguists to discover facts about language and the principles of teaching second languages within the CLT paradigm. As a result, linguists who were also language instructors began experimenting with using corpora directly with language learners in the classroom to facilitate language acquisition.

Corpora had been indirectly contributing to language instruction through their use in the creation of reference materials and textbooks for some time before Tim Johns, one of the first advocates for giving language learners direct access to corpus data, began criticizing these materials for keeping learners a step removed from the data. Johns felt that learners could benefit more from corpora by becoming language researchers themselves and analyzing the language data from a corpus first hand, a technique he named “Data-driven learning” (DDL).

Data-driven learning (DDL) is the use in the classroom of computer-generated concordances to help students explore patterns in the target language, and the creation of activities and exercises based on concordance output. Over the years, enthusiasm for using DDL and concordance output in the classroom with language learners has grown. Today, although there are a limited number of empirical studies outlining a clear connection between DDL and improved learning outcomes, applied linguists have outlined multiple theoretical reasons for using DDL with language learners. The main argument being that DDL creates learning conditions which have been found to facilitate second language acquisition (SLA) processes.

Vocabulary instruction is one area of language teaching currently being greatly informed by descriptive analyses of corpora, and which may have a growing need for DDL activities. Specifically, the creation of general academic and discipline specific wordlists from corpora are beginning to inform vocabulary instruction for English for Academic purposes (EAP) and English for Specific purposes (ESP) courses. At the same time, vocabulary instruction within the CLT paradigm is moving away from teaching words in isolation, and placing a greater emphasis on exposing learners to lexical items in authentic and meaningful contexts. Furthermore, there is a growing amount of evidence that much of the English language is formulaic (i.e., stored and retrieved in the mind as chunks of language), which suggests that teaching vocabulary as separate from grammar has limitations. DDL has been viewed as a possible technique which can keep vocabulary instruction current with the research by placing words from wordlists back into authentic and meaningful contexts. Furthermore, DDL, because it utilizes concordance lines of naturally occurring language, will contain many instances of formulaic language, which can be highlighted to a greater or lesser extent, to promote noticing of these linguistic structures.

In addition to its relevance to current developments in applied linguistics, DDL also has empirical support for its use in vocabulary teaching. Viewing concordance lines has been found to lead to small but consistent gains in students' vocabulary knowledge, greater recall, and to the acquisition of transferable word knowledge compared with traditional vocabulary teaching methods, such as using a dictionary.

While pitting DDL against traditional methods for teaching vocabulary is effective in proving that DDL is a powerful teaching technique, it is problematic in that it

seems to suggest that DDL activities should be favored over these methods and materials. Because it is unlikely that traditional reference materials and methods for teaching vocabulary will be abandoned by the language teaching profession, it is more realistic, beneficial and progressive to consider how to best exploit DDL in conjunction with more traditional methods of vocabulary instruction.

Although research has looked into the positive learning outcomes which are associated with DDL, there has been very little research which seeks to outline the unique contributions DDL may make to vocabulary teaching and learning. In other words, what do language learners learn about a word from DDL activities that they don't from more traditional activities? Furthermore, no research has looked into the advantages of using DDL over vocabulary learning strategies, such as guessing the meaning of a word from its context.

The goal of the current study was to examine students' perceptions of the benefits of one type of DDL activity over their own methods for studying unknown words in an attempt to begin to outline how and for what purposes DDL can best be exploited to teach vocabulary. In addition, because DDL has been proposed as a viable technique for teaching formulaic language (e.g., collocations), two designs of the activity were compared in order to establish if design plays a role in students noticing of these structures. Finally, because the activity used in this study was created by the instructor/research (though it was only amended slightly from an activity proposed by Sinclair [2003]), students' perceptions of the difficulty level of the current activity and its different steps were assessed in order to offer suggestions for improvement and caveats for future use of the activity. No specific hypotheses were formulated before the study

because it was intended to be explorative, as conclusions were based on a qualitative analysis of the students' comments. However, in the case of determining which design of the activity was better at promoting students' noticing of collocations, it was hypothesized that when concordance lines were sorted according to the collocation, as opposed to randomly, more students would notice them. The efficacy of this design was thought to be greater because it essentially eliminates a step for the student by categorizing the data according to the collocate rather than causing the student to have to categorize the data for this feature his- or herself, as is necessary in the group receiving the randomly sorted concordance lines.

The following chapter surveys the literature on corpus linguistics with a focus on its contributions to English Language Teaching (ELT). Then, the uses of DDL for vocabulary teaching will be discussed. Finally, the literature on formulaic language will be reviewed. Chapter three will describe the study, its methods, subjects, materials, procedures, and results. Finally, chapter four will discuss the results of the study with a focus on their implications for using DDL to teach vocabulary and collocations, and improving the DDL activity used in this study. In addition, directions for future research will be suggested.

CHAPTER TWO

LITERATURE REVIEW

2.1 History of corpus linguistics

Corpus Linguistics is a methodology within the field of linguistics that has been developing rapidly since the year 1964 when the first computerized corpus, The Brown Corpus¹, was completed. Corpus linguists are mainly interested in descriptive or functional interpretations of language (Meyer, 2002), and study linguistic phenomena through the empirical analysis of large computerized databases of language called corpora (corpus, sing.). A corpus is “a large and principled collection of natural texts” (Biber, Conrad, & Reppen, 1998, p. 4), which is compiled so that it is representative of the language in general, a dialect, or other subset of the language. Corpora may contain language based on written texts, transcribed speech, or both. These texts are stored electronically, and then analyzed using computer software programs called *concordance generators, concordancers*, or, generically, *concordancing software* (Conrad, 2005; Tribble & Jones, 1990).

Collecting large amounts of text in order to analyze linguistic phenomena was not a new concept when corpus linguistics arrived as a methodology. As Meyer (2009) points out in a recent article, early dictionaries were based on a large body of published works and millions of citation slips of naturally occurring language. Furthermore, concordance lines (i.e., a word displayed within a surrounding context) as a format for displaying

¹ Henry Kucera and W. Nelson Francis compiled the first computerized corpus, The Brown Corpus, at Brown University, which was completed in 1964. The size of the corpus was one million words; a large corpus for its time (Leech, 1997).

every instance of a word in a text or collection of texts has been around for centuries, as Tribble and Jones (1990) explain:

In its original sense a concordance is a reference book containing all the words used in a particular text or in the works of a particular author (except, usually, the very common grammatical words such as articles and prepositions), together with a list of contexts in which each word occurs. [...] Books like this have been in use since the Middle Ages, especially in Biblical Scholarship. [...] The earliest known complete concordance of the Latin Bible was compiled by the Benedictine Hugo de San Charo in the thirteenth century. Hugo, it is said, was assisted by no fewer than 500 monks. (p. 7)

In modern linguistics, the work of collecting and analyzing large databases of language, though still a time-consuming and tedious task, has been greatly simplified and largely automated by powerful computers and concordancing software programs, of which Laurence Anthony's *AntConc* (2008) is just one example. These powerful software programs have the capacity to rapidly and accurately locate every word in a text together with its surrounding context, including even those very common words like articles and prepositions that were a burden, and often not included, in concordances that were collected manually, such as those by the monks of Hugo de San Charo. In addition to their capacity to search for keywords and their surrounding contexts, these programs also have the capability of calculating frequency information about words, which is often presented in the form of hierarchical lists (usually with the most frequently occurring word appearing at the top of the list). Computerized corpus searches are not limited to word-level searches, however. Users may also choose to run searches on two or more words (i.e., collocations), phrases, clauses, or, if the corpus is tagged, on grammatical categories (e.g, prepositions and articles).

As this short history of corpora and concordancers shows, this method of studying language was not invented by corpus linguists nor was it incidental to the creation of the computer. Instead, old methods of doing linguistic research have been greatly supported by modern technology, and this has caused linguists to return to the “empirical tradition”, to borrow a term from Meyer (2009). A tradition which Meyer claims, “fell into disfavor following the rise of Chomsky’s theory of generative grammar in the 1960s” (p. 208). This return to the empirical tradition in linguistics comes at the same time that language teaching theorists are emphasizing the facilitative effects of exposing learners to authentic examples of language in the classroom; teachers in addition to their counterparts in research are beginning to find uses for corpora.

2.2 Corpora in linguistic research and language teaching

Since their beginning, computerized corpora have been mainly used for research or “for finding out about language and texts” (Leech, 1997, p. 2). Today, nearly every subdiscipline within linguistics uses corpora, to a greater or lesser extent, to inform their studies.

Although corpora have been used by linguists for research purposes for over forty years, researchers who are also language instructors are beginning to have greater interest in exploiting corpora for the teaching of second and foreign languages. According to Leech (1997), corpora can have a direct or indirect effect on the language classroom. Indirectly, corpora are impacting the language classroom because they are being used by materials developers to create improved reference materials (e.g., dictionaries, grammars, and thesauri) and textbooks. Furthermore, corpora are being exploited by language instructors to inform syllabus and course design (Flowerdew, 1993), and to create tests

(Coniam, 1994; Shillaw, 1994). Moreover, corpora have been used to create both general academic (Coxhead, 2000, 2002), and discipline specific (Wang, Liang, & Guang-chun, 2008) wordlists. Wordlists like Coxhead's (2000) Academic World List (AWL) contain the most frequently occurring headwords of a discourse; in the case of the AWL, the words are those which occur most frequently in general academic discourse, regardless of discipline. Coxhead's list is based on three principles: teaching the most relevant, useful, and frequent lexical items to students first. The list has contributed to the prioritization of vocabulary for the EAP curriculum. However, while useful for prioritizing vocabulary instruction, wordlists need to be taught using a principled approach to teaching vocabulary accompanied by appropriate classroom techniques in order to assure that students acquire and are able to correctly and creatively use these words in their own speech and writing. Corpus-based methods and activities can help. This brings us to the discussion of how corpora are having a direct effect on the language classroom.

Essentially, there are two ways to directly engage second language learners in corpus work in the classroom: 1) they can be given direct access to a corpus and concordancing program on a computer; or 2) they can be given print-outs containing the raw data, or concordance output, from a corpus. The discussion here will be restricted to the latter option as the current study focused on teaching academic vocabulary to learners who had never engaged in corpus-based work prior to the study, and therefore needed to be exposed to this "first stage" of corpus consultation in order to become familiar with how to use and analyze concordance results (Leech, 1997; Chambers, 2007); though, students should ultimately be given access and taught how to use online corpora to encourage and support autonomous language learning beyond the classroom.

The term “Data-driven learning” (DDL) has been coined to denote activities in which language learners are given printouts of computer-generated concordances in the classroom in order to explore language patterns. The term was coined by Tim Johns to describe a method he used, and was largely responsible for developing and popularizing (Johns, 2000a, 2000b). Johns’ DDL is largely based on the methods used by the linguists involved in the COBUILD project at Birmingham University. This project, directed by John Sinclair, made extensive use of key-word-in-context (KWIC) concordance data in order to create a range of reference and teaching materials for English language learners, most notably dictionaries. The project was revolutionary in that it was the first attempt by lexicographers to build a comprehensive profile for each word entry in a dictionary based on empirical evidence of actual native speaker use of the language. Inspired by the project’s approach toward the description of linguistic phenomenon, Johns perceived the benefit this new technology and methods of analysis could bring to the language learner. Therefore, he developed DDL in order to “cut out the middleman as far as possible and to give the learner direct access to the [language] data” (p. 30). Johns (2002) concisely describes the difference between the goals of his DDL and those of the COBUILD project in the following excerpt:

From the start, it was clear that there would be a small but significant difference between the approach taken by our colleagues in COBUILD and what I wanted to do in the English for International Students Unit. In the COBUILD materials the data was to some extent ‘hidden’ from the learner by the team of researcher [sic] and lexicographers. My approach was rather to confront the learner as directly as possible with the data, and to make the learner a linguistic researcher. The metaphor I use with my students is that of the detective. (pp. 107-108)

Over the years, a significant body of research has developed which promotes the use of corpus-based activities, such as DDL, with second language learners. Many strong theoretical arguments have been made which align corpus work with principles of second-language acquisition (SLA) and situate it within the communicative language teaching (CLT) paradigm. Researchers who have commented on the potential facilitative effects on SLA processes of corpus consultation by second language learners believe that the same procedures which corpus linguists use to conduct descriptive studies of language can be taught to and used by language learners themselves to promote SLA processes. Using corpus analysis methods with second language learners has the added effect of reassigning traditional classroom roles, whereby students become linguistic “researchers”, and teachers become directors or coordinators of research (Gavioli, 1997, 2001; Johns, 2000a, 2000b). However, Bernardini (2002) states that “descriptive insights and research methodologies have not simply been borrowed from the descriptive paradigm, but have been adapted, reformulated, and often extended in various ways to fit pedagogic concerns and priorities” (p. 29). In this newly structured classroom, students are encouraged to engage in linguistic research which involves raising questions about the target language and engaging in a process of hypothesis formation and testing of particular rules of the language, a process by which interlanguage development is thought to progress (Conrad, 2005). Furthermore, Aston (1995, 1997) argues extensively that “corpora can play a useful role in the acquisition and restructuring of schematic knowledge” (1995, p. 263). His argument is that concordance lines expose learners to contextual repetition and variation of linguistic structures, promoting a process of synthesis and analysis of information on the part of the learners, which, in turn is a key to

acquisition. Finally, many researchers have noted that engaging students in corpus-based activities promotes noticing or consciousness-raising (Kettemann, 1995; Johns 1991a, 1991b; Conrad, 2005).

In addition to facilitating SLA, corpus-based activities are viewed as being consistent with a variety of principles and learning goals within the CLT paradigm, which currently dominates the English language teaching (ELT) profession. First, concordance output exposes learners to linguistic phenomena in authentic contexts, which learners have to analyze and categorize inductively (i.e., they must categorize the data and are thus lead to discover the rules of the language on their own). Furthermore, the redefined role of the learner as researcher shifts control of learning from teacher to student, causing the classroom to become more student-centered during these activities. Finally, corpus-based activities are thought to increase learner autonomy “as students are taught how to observe language and make generalizations rather than depending on a teacher” (Conrad, 2005, p. 402).

In addition to the benefits of using corpora in the classroom, there also exist many caveats which have to be considered before engaging learners in this type of work; the most important being the issue of training learners to analyze and categorize linguistic data. Gavioli (2001) explains that training learners to work with corpus data is difficult because “Unlike dictionaries, grammars and textbooks [concordance data] does not offer explanations; it merely provides data which it is up to the user to explain” (p. 110). The difficulty of training nonnative speakers to analyze and make generalizations about the target language is further compounded by the fact that they are unable to rely on their intuitions to help guide their analyses as native speakers were found to do in a 2004 study

by Sripicharn. In addition to making the process harder, a lack of intuition about the target language often leads nonnative speakers to make overgeneralizations about the language. However, these same nonnative-speaking students were also found to adopt alternative strategies, such as forming and testing hypotheses, since they could not rely on their intuitions, suggesting that with additional guidance from the language instructor, nonnative speakers can benefit greatly from analyzing concordance output.

2.3 Data-driven learning (DDL) and vocabulary instruction

Because corpus-based activities, such as DDL, have strong theoretical reasons backing their implementation in the classroom with language learners and because these activities are viewed as being in alignment with current language teaching philosophies, there has been a growing number of publications outlining a wide range of uses for DDL in the classroom of which teaching vocabulary is just one.

Many researchers have outlined the uses and benefits of using DDL or concordance output to teach vocabulary to second language learners (Cobb, 1997, 1999, 2007; Horst, Cobb, & Nicolae, 2005; Pickard, 1994; Stevens, 1991a, 1991b; Thurstun & Candlin, 1998). These studies have demonstrated how to create more traditional vocabulary activities (e.g., fill-in-the-blank or matching exercises) that have the added advantage of being based on authentic texts and also exposing learners to multiple, novel contexts at one time (Stevens, 1991; Thurstun & Candlin, 1998). Researchers have also created and demonstrated the efficacy of complex online self-access vocabulary packages for extensive vocabulary study using concordance data alongside more traditional reference materials (Pickard, 1994; Horst, Cobb, & Nicolae, 2005).

In addition to supplying teachers with ideas for creating corpus-based vocabulary activities, researchers have outlined the facilitative effects of using corpus-based materials. For example, Stevens (1991) found that concordance-based vocabulary exercises can be more easily solved by learners than traditional gap-filler exercises, suggesting they should be used “if the purpose of the exercise is to reinforce the vocabulary, as opposed to testing, and if the proclivity of the teacher is to engender a sense of confidence and well-being in the students” (p. 55). However, Stevens could not make any claims about the efficacy of the activity because he did not empirically test learning outcomes from engaging in DDL. Tom Cobb’s (1997, 1999) work is the only research which has empirically tested the effectiveness of corpus-based techniques to teach vocabulary. Picking up where Steven’s study left off, Cobb (1997) compared the vocabulary learning outcomes of his students when new words were learned by viewing multiple concordance lines vs. a single sentence accompanied by a short definition of the word, and found that viewing concordance lines lead to small but consistent gains in his students’ vocabulary knowledge. Furthermore, in a follow up study, Cobb (1999) found that viewing concordance lines also facilitates the acquisition of transferable word knowledge, supported by the fact that these students were able to apply their knowledge of the word in novel activities and contexts.

This empirical evidence suggests that DDL has an important and meaningful place in the vocabulary teaching curriculum. However, many studies have tended to pit DDL against other more traditional vocabulary teaching activities and materials in an effort to see which leads to greater learning outcomes. Although doing this proves that DDL is an effective and powerful learning tool, it is problematic because it seems to

suggest that DDL should be favored over these materials which have been facilitating language acquisition for centuries.

Instead of promoting DDL over the use of more traditional vocabulary teaching methods and activities, DDL should be seen as but one part of a holistic plan to teach vocabulary and used in conjunction with traditional methods and activities. However, because DDL has often been viewed as in competition with traditional methods for teaching vocabulary, there has been little research on the unique contributions this activity has to make to the teaching of vocabulary.

Therefore, the main goal of the current study was to determine how DDL can be used to complement traditional vocabulary learning methods (i.e., with dictionaries and by guessing the meaning of a word from its context). In other words, what can DDL teach students about a word that traditional methods either cannot or fail to? A second goal of the study was to determine how best to design offline DDL activities in order to support students' analyses and to teach collocations. Since much of the research has either focused on teaching vocabulary with online concordancing (Cobb, 1997, 1999) or by using concordance lines to amend the format of traditional vocabulary exercises (e.g., gap-fillers) (Stevens, 1991), there is a need for research which seeks to discover how to best design and exploit DDL activities in their purest, offline form (i.e., as raw language data extracted from an appropriate corpus and subsequently given to language learners for analysis). Furthermore, this study contributes to the growing body of literature which is seeking to discover how the theoretical benefits of corpora are being realized through the collection of students own accounts of their experiences working with concordance data (Boulton, 2008; Chambers, 2005; Götz & Mukherjee, 2006).

Before the study and its results are discussed, however, a review of the literature on formulaic language follows, as this is a fairly new and important area of current linguistic enquiry that has implications for teaching vocabulary to nonnative speakers, and for this study, which examined the efficacy of two DDL designs in promoting the noticing of collocations – one type of formulaic language.

2.4 Formulaic language

In her 2000 article “Formulaic Sequences in Second Language Teaching: Principle and Practice”, Alison Wray outlines how linguists have begun to distinguish formulaic from non-formulaic language. One of the most prevalent methods for observing formulaic language has been through systematic analyses of large representative corpora. Large corpora are searched for evidence of formulaic language through either raw frequency counts or statistical measures. In addition to these methods, formulaic language can be identified in: 1) The phonological form of words and phrases: children tend to enunciate less clearly chunks of language which they have learned whole and for which they have not yet acquired control of the constituent parts, and it could be argued that consonant weakening and vowel reduction found in adult native speakers’ continuous speech constitutes a formulaic sequence; 2) Code-switching: when bilinguals code-switch, it is thought that they do so at the boundaries of formulaic sequences; 3) Interlanguage phenomena: fossilized errors or the correct use of a construction otherwise not within the scope of the speaker’s grammatical competence can indicate that the sequence has not been created from scratch (Wray, 2007, p. 467).

Most relevant to language instructors interested in teaching formulaic language is the fact that the some of the same methods for locating formulaic language in corpora can

be used with language learners to help them acquire these linguistic structures. However, it is first important to know what these structures are, and how they are realized.

There are many definitions, names, and ways of categorizing formulaic language. However, the criterion for labeling a multi-word sequence as formulaic is usually dependent upon the frequency with which two or more words appear together. Furthermore, linguists (Bahns, 1993; Cantos & Sánchez, 2001; Lennon, 2005; Fan, 2009) also distinguish between types of formulaic sequences based on the transparency of meaning (i.e., if the meaning of the phrase can be understood from its constituent parts) of the string of words, and if they are psychologically salient. These linguists place language fixedness on a continuum which is represented by three categories: free combinations, collocations, and idioms. Free combinations or non-formulaic language are word combinations that can be formed by choosing lexical items based on their common meaning(s) and applying grammatical rules to combine them, while idioms are “a group of words that occur in a more or less fixed phrase and whose overall meaning cannot be predicted by analyzing the meanings of its constituent parts” (Simpson & Mendis, 2003, p. 423). However, some linguists argue that the meanings of idioms may be more or less transparent and thus analyzable from its constituent parts (Fan, 2009). Collocations occupy the middle ground between completely fixed combinations such as idioms and free combinations. However, Bahns (1993) warns that there are “‘transitional areas’ between free combinations and collocations and collocations and idioms” (p. 57). Indeed, collocations seem to be the most elusive of the three categories to researchers, instructors, and second language learners alike. Therefore, they warrant closer analysis if they are to be taught to language learners using DDL activities.

According to Bahns (1993), “The main characteristics of collocations are that their meanings reflect the meaning of the constituent parts (in contrast to idioms) and that they are used frequently, spring to mind readily, and are psychologically salient (in contrast to free combinations)” (p. 57). Bahns further categorizes collocations by placing them into two categories: lexical collocations and grammatical collocations. Grammatical collocations (or colligations) “consist of a noun, an adjective, or a verb, plus a preposition or a grammatical structure such as an infinitive or clause [and examples] include *account for, advantage over, by accident*” et cetera, while lexical collocations “do not contain prepositions, infinitives, or clauses, but consist of various combinations of nouns, adjectives, verbs, and adverbs” (p. 57).

Both types of collocations referred to above are *structural* collocations. These are just one type of collocation. Fan (2009) discusses two additional categories of collocations: 1) semantic collocations: collocations where the semantic properties of the item(s) determine its collocates; and 2) statistical collocations: the occurrence of two or more words within a short space of each other (in this type, there are frequent and non-frequent collocations, and no collocation is impossible).

2.4.1 Reasons for teaching formulaic language

One reason for including formulaic language, such as collocations and idioms, in the language teaching curriculum is that they exist in large quantity in the English language, regardless of register, and perform a variety of communicative functions. For example, Simpson and Mendis’ (2003) study of idioms in academic speech in the MICASE corpus lead them to conclude that “The discovery of a significant number of idioms in a corpus of academic speech, and, more importantly, the evidence that they

perform a variety of important pragmatic functions provides the rationale for including them in an EAP curriculum” (p. 432).

Other researchers (Ellis, Simpson-Vlach, & Maynard, 2008) point out that the second language learner has to know formulaic language in order to increase their reading speed and comprehension, and to be able to write in a native-like fashion. In fact, many researchers (Simpson & Mendis, 2003; Wray, 2000; Fan, 2009; Siyanova & Schmitt, 2008) agree that learning formulaic language enables learners to sound or write in a more native-like fashion. Furthermore, Simpson and Mendis (2003) state that language learners tend to share this point of view which often motivates them or translates into a desire to acquire as many idioms or language chunks as possible (p. 419). However, instructors should be careful not to assume that their learners are concerned with sounding native-like, and be conscious of this fact before requiring students to have a strong command of idioms and collocations.

Another important reason to teach second language learners formulaic language is discussed by Bahns (1993), which found that “learners’ knowledge of collocations lags far behind their knowledge of vocabulary in general” (p. 101). Bahns further states that one reason non-native speakers’ lack receptive and productive knowledge of collocations is because of vocabulary teaching methodology. Brown (1974) suggests a remedy: “if we [instructors] choose instead the collocational group as the practice unit [as opposed to the lexical item combined with grammar rules] we are reinforcing the fact that we [native speakers] both read and speak in ‘chunks’ of language giving flexible variety, and letting the students get the ‘feel’ of useful groups of words” (p. 2). One way to shift the teaching focus to language chunks is by using DDL to highlight these structures.

2.4.2 Difficulties of teaching formulaic language

Some scholars fear, however, that collocations are too numerous to be taught and that the only way non-native speakers are to acquire some degree of competence in using them is through years of study, reading, and observation of the language (Bahns, 1993). Indeed, learners face many barriers when attempting to acquire formulaic language. However, most of the difficulties students will face in learning formulaic language are also difficulties they face when attempting to acquire the lexis and grammar of a language in general. For example, students' first language background plays an important role in determining the difficulty a student will have learning new formulaic sequences. Also, lack of exposure to the target language or comprehensible input can be a factor. Other factors include: the rare occurrence of some collocations and idioms, the fact that many idioms do not appear in their canonical form in academic discourse (e.g., many are subject to truncation, creative blending, or performance variations), rendering most treatments of idioms and collocations in special dictionaries and course books ineffectual because students are unlikely to meet these formulaic sequences in the form that they appear in these materials. Furthermore, idioms and some collocations rely on and assume a specific cultural schema for interpretation which is difficult for learners to understand and interpret.

Keeping the reasons for teaching formulaic language and the difficulties learners face in learning it in mind, many researchers have suggested methodologies and techniques for teaching formulaic language (Brown, 1974; Ellis, Simpson-Vlach, & Maynard, 2008; Farghal & Obiedat, 1995; Gardner & Davies, 2007; Kennedy, 2003; Lennon, 2005; Man-lai, Pui-yui, & Chau-ping, 1994; Siyanova & Schmitt, 2008; Sun &

Wang, 2003). Many of these methodologies and techniques echo the principles outlined in CLT; namely, that formulaic language should be taught using authentic texts, both explicit and implicit teaching strategies, and students should learn through repeated exposure and authentic production.

If the best method for teaching second language learners formulaic language is through multiple exposures to formulaic sequences in authentic, and preferably meaningful, contexts, then DDL seems especially appropriate for bridging theory and practice, though few studies on collocations mention DDL as a solution for teaching them. This study fills this gap by attempting to use DDL to teach collocations to university students.

CHAPTER THREE

THE STUDY

This first section of this chapter will discuss the context within which this study took place and the materials and methods used to answer the following research questions:

- 1) What are the students' perceptions of the difficulty level of a DDL activity designed to teach the parts of speech, meanings, and collocations of the academic vocabulary word *subsequent*?
- 2) What do students think they learned about the word *subsequent* from viewing it in multiple concordance lines that they think they would not have if they attempted to learn the word using their usual method(s) for learning new vocabulary words?
- 3) What are the students' perceptions of what they are being taught or learning by completing the DDL activity?
- 4) In which design do more students notice the collocate *to* of the target word *subsequent*?: When concordance lines are sorted according to the collocate of the target word (Group 1) or when they are sorted randomly (Group 2)?

These research questions were asked in order to satisfy a larger goal of determining how English language instructors can best design and exploit DDL activities to teach academic vocabulary. Furthermore, research question two was asked in an effort to isolate the unique knowledge the students' perceived they acquired from DDL in order to outline where DDL best fits into the language teaching curriculum as a technique to teach unknown vocabulary alongside more traditional methods. In other words, what does DDL teach students about a word that no other resource (e.g., a dictionary) does?

Following the discussion of the context, materials and methods of the study in section one, the results of the study will be shared in section two.

3.1 Methods

This study used a semi-structured interview format to elicit students' perceptions of a DDL activity immediately following their completion of the activity. In addition, the students' background information, beliefs of and attitudes towards grammar and vocabulary, and methods for studying unknown words were collected through questionnaires.

3.1.1 Subjects and context

Fifteen, ESL students enrolled in an EAP course at a large, U.S. university took part in this study. The class was extremely heterogeneous, with students from eight different nations: Japan, Iran, Kuwait, Saudi Arabia, Pakistan, Liberia, Nigeria, and South Korea. Furthermore, the students spoke six different native languages: Japanese, Farsi, Arabic, Kru, Yoruba, and Korean. A more detailed description of the students' background, including the number of years studying English, amount of time in the U.S., and age, can be found in Appendix A. All the students consented to take part in the study and were given pseudonyms to assure that their identity remained anonymous. Both the training sessions and the experiment took place during hours of regular instruction.

3.1.2 Materials

In order to create the materials for the training sessions and the experiment, the researcher/instructor compiled a corpus composed of texts taken from the students' assigned textbook, *Down to Earth Sociology: Introductory Readings* by James M. Henslin, for the course. A detailed word count for the corpus can be found in Table 1 below.

No. of texts	Title of Text	Token Total
1	The Rules for Giving Christmas Gifts	3,782
2	Doing Fieldwork among the Yanomamö	7,926
3	The Sound of Silence	4,000
4	Caught Between the Ages	5,446
5	Eating Your Friends is the Hardest	4,530
6	Would you Hire an Ex-Convict?	3,086
7	If Hitler Asked you to Electrocute a Stranger, Would you? Probably	3,139
8	Attacking Nicely: Women Selling Cars	5,948
9	But What do you Mean? Women and Men in Conversation	2,583
10	The Importance of Being Beautiful	2,869
	Total Running Words in Corpus	43,309

Table 1: Corpus of G011 Sociology Texts Word Count

The corpus was created by scanning each assigned text from the course textbook to a computer, and then reading each text with optical character recognition (OCR) software in order to convert the scanned file to a text file. Before each text file was saved, it was manually checked for any errors. OCR software is a convenient and largely accurate way to avoid manually typing each text into the computer, but the software is not perfect, and more than a few errors will usually occur in each text. Therefore, it is necessary to read the full texts manually before they are used for creating DDL activities. After the text files were cleaned up and saved, they were searched using Laurence Anthony's *AntConc version 3.2.1* concordancing software. Though there are many concordancers available on the market, Laurence Anthony's software was chosen because it is a free download, available on his webpage at <http://www.antlab.sci.waseda.ac.jp/software.html>, and contained all the search capabilities needed to create DDL activities for this study.

In addition to using the corpus to create the DDL activities, the corpus was used to prioritize vocabulary instruction based on the frequency and distribution of the words

in these texts, in addition to other criteria. The words used in this course and study were chosen by first restricting the choice of words to those 300 academic words which appear in the first five sub-lists of Coxhead's (2000) AWL. The first five sub-lists of the AWL were chosen because the words in these lists were found to account for 8.3% of the words in Coxhead's corpus, whereas sub-lists six through ten only had 1.7% coverage.

Therefore, sub-list five is thought to be the cutoff point for the best return for learning (Coxhead, p. 76). Next, only those words from the AWL which occurred frequently (i.e., five or more times including inflectional forms) in the Corpus of G011 Sociology Texts or that were thought to be of general importance (i.e., serve a strong discourse function in academic speaking/writing) were included in the list of vocabulary items to be learned over the semester (16 weeks). This resulted in a total of seventy-three headwords that the students were to learn over the course of the semester. A complete list of these words can be found in Appendix B. In addition, the order in which each word was introduced to the students depended on which text the word appeared most frequently in so that the students were learning the vocabulary as they encountered them in their reading assignments.

Because the Corpus of G011 Sociology Texts was small, search results did not return many instances of the words from the AWL. In addition, because the corpus was unbalanced (i.e., contained texts from the sociology field only), search results could not be considered representative of academic language in general. Therefore, a larger corpus, Mark Davies' Corpus of Contemporary American English (COCA) (academic sub-section only²), was additionally consulted in order to obtain more instances of the word and more representative results. COCA was also an obvious choice for a corpus as it is

² The academic section of the COCA consists of 76 million words.

freely available on the web at www.americancorpus.org, contains its own built in concordancer, SARA, and is composed of the type of language the students were concerned with acquiring, namely American, academic English..

3.1.3 Training sessions

Training sessions for students who have never worked with corpus-based activities is a necessary prerequisite to conducting experiments on student use of, attitudes towards, and/or learning outcomes of DDL activities (Conrad, 2005; Gavioli, 1997; Götz & Mukherjee, 2006; Stevens, 1991; Tribble & Jones, 1990). No student who took part in this experiment had heard of a corpus or engaged in corpus-based activities before this experiment. However, the researcher was encouraged to introduce data-driven learning coursework to her students by studies that have reported on the success of engaging first-year, college-level students in this type of work (Chambers, 2005; Hadley, 2002).

In order to prepare students for DDL work, they participated in five, 45-minute training sessions. In session one, the students were given a brief introduction to corpus linguistics. Only those items that would help the students understand and appreciate the DDL materials were discussed since these were English language learners, not students of linguistics. Specifically, the students were told what a corpus was, what concordance lines were, and they were shown how their activities were created by performing an in-class search of the word *consist* using COCA. This brief introduction was given so that the students would understand and not be distracted by the truncated concordance lines in the DDL activity, and so they would better appreciate the information (e.g., grammatical features, collocates, discourse functions, etc.) about a vocabulary item that is revealed by

performing computerized searches of a keyword, and which is subsequently represented in printed form when it is given to them as a DDL activity.

In session two, the students were given a short lecture accompanied by a hand-out on collocations and idioms so they would understand how to approach the section of the DDL activity that asked them to identify these language chunks in the data. The hand-out can be found in Appendix C. It describes and provides examples of the different levels (i.e., collocations and idioms as distinguished from free combinations) and types (i.e., lexical and grammatical) of formulaic language.

Following these introductions to corpora and formulaic language, the students engaged in DDL work for three sessions. In these sessions, the students completed a DDL activity which asked them to answer a series of questions about a target word by analyzing and categorizing concordance output extracted from the Corpus of G011 Sociology Texts and COCA. The design of the activities was borrowed, and amended slightly, from John Sinclair's (2003) *Reading Concordances*. An activity from one of the training sessions can be found in Appendix D.

In these activities, the students were asked to complete six steps. First, the students separated the concordance lines according to the target word's part of speech. Then, based on the part of speech of the word, they categorized the concordance lines again; this time, according to their perception of the meaning(s) of the target word. In this step, students were asked to think of the closest synonym(s) of the word and to categorize all instances of a meaning under a corresponding synonym. Next, the students were to simply identify the number of senses of the word they had found; after which, they again analyzed the concordance lines for the meaning(s) of the word in order to make sure that

they were not making distinctions where there were none, and conversely, that they didn't place two unrelated senses under the same synonym. After this second round of analysis, the students wrote a brief definition for each sense of the word. Although the definition was to be brief, it needed to be longer than one word, and synonyms were only acceptable if they were accompanied by a longer description of the meaning of the word. Next, the students were asked to look for collocational patterns and idioms in the language data, and to note any they found. As a final step, the students created their own sentences for each meaning of the target word.

3.1.4 The experiment

At the beginning of the semester, the students completed a questionnaire about their attitudes towards grammar and vocabulary and their common method(s) for studying unknown vocabulary. The questionnaire can be found in Appendix E. In addition, the students were randomly placed into two groups: Group 1 (N=8) and Group 2 (N=7). Both groups completed a thirty-item vocabulary test on a random selection of six words from each of the first five sub-lists of the AWL. The format of the test questions can be found in Figure 1 below. The format of the test is based on a design proposed by Wesche and Paribakht (see Nation, 2001), and was chosen based on its sensitivity in accurately identifying students' knowledge of a given word while also eliminating the chance that correct responses may be the result of a lucky guess, which can occur in traditional cloze or matching vocabulary tests. The test was administered to the students in order to identify a word which no student had prior knowledge of and that would subsequently be used for the DDL activity in the experimental session. Furthermore, identifying a word which was unknown by every student in the study made it possible to

establish a baseline of the students' knowledge in order to understand and compare the results of the experiment.

<p>Subsequent</p> <p>a. I don't remember having seen this word before.</p> <p>b. I have seen this word before, but I don't know what it means.</p> <p>c. I have seen this word before, and I think it means _____ (synonym or phrase)</p> <p>d. I know this word. It means _____ (synonym or phrase)</p> <p>e. I can use this word in a sentence (If you do this section, please also do letter d.):</p>

Figure 1: Sample Question from AWL Vocabulary Pre-Test

From the results of this test, the word *subsequent* was chosen for the experiment session. Most students (N=11) answered either letter *a*, "I don't remember having seen this word before", or letter *b*, "I have seen this word before, but I don't know what it means". Of the remaining students (N=4), two answered with letter *c*, "I have seen this word before, and I think it means _____", one answered letter *d*, "I know this word. It means _____", and one answered letter *e*, "I can use this word in a sentence". However, two of these students' answers did not contain the correct meaning or use of the word. Of the two students remaining, one replied by supplying the synonym "sequential", and the other replied by writing "subsequently, he passed the exam". Though these responses are somewhat accurate, they do not reveal a strong understanding of the word. The first student uses a vague and difficult synonym, where simply the word *after* or the phrase *to come after* could have been used. The second student's response is grammatically correct, but failed to reveal much about the meaning of the word, showing that this student has a strong understanding of the English adverbial system, but not necessarily a strong understanding of the meaning or use of this word. Furthermore, both of these students admitted to not knowing the meaning of the word in the post-experiment interview. Thus, all of the students had little to no prior knowledge of the word *subsequent*, and the results

of the experiment reveal the students attitudes toward and difficulties with a DDL activity designed to teach an unknown word.

After establishing the target word to be used in the experiment session, a keyword search was performed for the word in the Corpus of G011 Sociology Texts and COCA in order to compile a language data set which included all the inflectional forms of the word and all the senses for each inflection. In addition, a search for frequently occurring collocates was conducted. After the analysis was completed, fifteen representative concordance lines³ were chosen for the activity. Using the same set of concordance lines for each activity, the lines were sorted either according to the word's most frequently occurring collocate (Group 1), or at random (Group 2) in order to answer research question number two (see above). The activities created for Groups 1 and 2 can be found in Appendices F and G respectively.

During one week of regular instruction, each group came to only one of their regularly scheduled class periods so that the groups completed the experiment sessions separately. For the experiment, the students were asked to complete a DDL task for the word *subsequent*, identical to those they completed during the training sessions. Each group of students was given forty-five minutes to complete the activity. They were not allowed to ask questions or to work with their classmates, and were instructed to complete the activity to the best of their abilities. After the activity session, each student was privately interviewed by the researcher/instructor using a semi-structured interview format. The post-activity, interview protocol can be found in Appendix H. The interviews

³ Fifteen concordance lines is a somewhat arbitrary number, but the amount of lines included was based on two criteria: including at least five examples of the same meaning, part of speech, and frequent collocations of the target word, and keeping the number of concordance lines at a minimum so that the activity could be completed within the allotted time period, forty-five minutes.

were video/audio recorded, transcribed, and analyzed for recurring patterns and salient features by the researcher/instructor.

3.2 Results

The results of the study will be discussed in the order of the research questions.

3.2.1 Results for research question one

RQ1: What are the students' perceptions of the difficulty level of a DDL activity designed to teach the parts of speech, meanings, and collocations of the academic vocabulary word *subsequent*?

The results of the students' responses to the first five questions of the post-treatment protocol (see Appendix H) are summarized in Figure 2 below. These questions asked the students to rate the level of difficulty they perceived they had in completing the entire DDL activity and each of its steps⁴, according to a three-point scale (easy, moderate, or difficult).

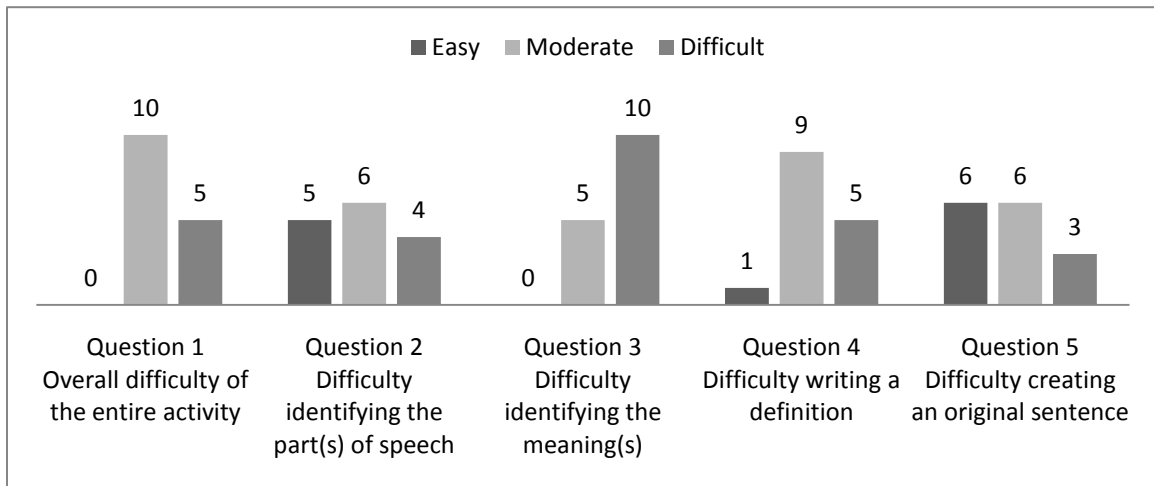


Figure 2: Student Perceptions of the Difficulty Level of the DDL Activity

The data collected for Question 1 as reported in Figure 2 above, reveal that none of the students rated the overall difficulty level of the activity as easy, 67% (N=10) reported

⁴ Data on students' perceptions of the difficulty of locating collocations was omitted because only 40% of the students completed this task.

it was moderately difficult to complete, and 33% (N=5) found the activity difficult to complete.

With regard to the discrete tasks, creating an original sentence was more frequently reported as easy (N=6 or 40%) by the students than any other task, while identifying the part(s) of speech of the target word was reported as an easy task by the second largest number of students (N=5 or 33%). The tasks of identifying the meaning(s) of and writing a statement/definition about the word were largely viewed as either moderately hard (N=14 or 47%) or difficult (N=15 or 50%) to complete by the students, with the former task rated as difficult to complete by the largest number of students (N=10 or 67%).

3.2.2 Results for research question two

RQ2: What do students think they learned about the word *subsequent* from viewing it in multiple concordance lines that they think they would not have if they attempted to learn the word utilizing their usual method(s) for learning new vocabulary words?

Prior to the experiment session, the students were asked about their strategies for learning unknown vocabulary (see Appendix E). In total, seven strategies were named by the students: 1) using the dictionary (print and/or electronic), 2) guessing the meaning of the word from its surrounding context, 3) memorizing the meaning of the word, 4) asking someone about the word, 5) using the word in real life, 6) analyzing the parts of the word to arrive at its meaning, and 7) taking notes about the word. The majority (80%) of the students named using the dictionary (33%), guessing the meaning of the word from the context in which it appears (20%), or both (27%) as their strategy for teaching themselves new English words before enrolling in this course. Therefore, an analysis of the data was restricted to the responses of those students who used these strategies to teach themselves

new vocabulary, as they were used by the majority of the students. Furthermore, those students who reported using other strategies to learn unknown vocabulary often used one or both of these strategies in addition to their other strategies to learn new words, and those who did not were too few to be able to establish clear themes of the differences between the two learning methods as perceived by the students.

Students who reported using the dictionary as their main strategy for acquiring an understanding of new words felt that the DDL activity helped them to notice the different forms, meanings, and uses of the word. In addition, they reported noticing a connection between a word's form and its meaning. Some students remarked that the knowledge they acquired about a word from DDL could either not be acquired from using a dictionary or that they were not in the habit of using the dictionary to teach themselves such knowledge (e.g. multiple meanings and forms of the word). These students' comments will be discussed further in the following paragraphs.

Some of the students who reported using the dictionary to teach themselves new vocabulary remarked that when using a dictionary they do not pay attention to the different inflections and/or derivations a word takes while during the DDL activity they did. These students' comments are outlined in Table 2 below.

Maram	“it teach me also uhm uhm (1) if I check a word in dictionary I just understand that it is verb I don’t understand that for I don’t check up or down of that uhm word that understand that what is the adjective of that word or noun or adverb but here I can understand that it has how many + how many branches that one word has”
Kamal	“today’s activity is different [than the dictionary] + like + uhm it helps us to understand like the speech of part”
Hanok	“it’s helped me with to know like not just for example it’s just a verb because I know it like verb or the noun like many kind of part of speech”

Table 2: Student Comments on Learning Parts of Speech in the Dictionary vs. DDL

In addition to noticing the different inflections and derivations of the word, these students feel the activity gives them a better understanding of the relationship between the part of speech and the meaning and use of the word. A selection of the students’ comments on this topic can be found in Table 3.

Talal	“[in this activity] you can understand like what part of speech is the word because if I c- look it up in the dictionary I might find a meaning but I don’t know what if it’s like for for or for noun [...] or adjective [...] because here you can understand the: word the meaning of the word and th- how it + it comes like if it can be noun and uhm with different meaning and it can be adjective with different meaning”
Kamal	“[today’s activity] it helps us to understand [...] how like we can find the meaning by the speech part + like sometimes we find a word + we can find this a verb + and if it’s a verb + it should mean it has something different from the noun maybe they do they have a completely different meaning”
Hanok	“ in the dictionary it’s crazy they give you like subsequent and they give you the adverb the adjective and this and that but these things are used in various sentences so that you can see okay in this sentence it’s used as a noun in this sentence it is used as an adverb in the sentences it is used as an adjective in this sentence it is used as a verb + now you’ll be able to figure it out okay right now I can tell okay this is a verb there is the way it can be used [...] there is different from that of the dictionary”

Table 3: Student Comments on the Connection between Word Forms and Meaning

Lastly, students who normally consult a dictionary to learn new words stated that the DDL activity: 1) highlighted the polysemy of words for them where when they use the dictionary they taught themselves only one meaning at a time, and 2) helped them to understand difficult words that they have a hard time understanding the meaning of from the dictionary alone. Furthermore, one student, Hanok, commented that viewing concordance lines helped him understand the different functions of the word because he could see how they were used by others while simultaneously condemning dictionaries because of the isolated way words are displayed in them, likening this isolated format to that of the dropping of a bomb. These students' comments can be found in Table 4 below.

Eri	“oh in this activity I can see a lot of meaning but in this activity [using the dictionary] I usually do uhm I look only one meaning”
Sultan	“I think this way [DDL] is better because we have to use some difficult words and sometimes when I read the textbook or something I have a hard time figure any of the words even if I use my dictionary Arabic dictionary”
Hanok	“the dictionary will just + just drop the word BOMB and then give the definition but this + look at this activity you have you have fifteen different ways in which the word is used [...] now it give you broader idea of what the meaning or what the word should be [...] and how I should be able to use it in a sentence” [...] “I think the bottom line is uhm I realize that words can be different the dictionary will give the word but that particular word somebody can use in different way [...] you will not just see the word and just say there is the exact meaning unless it is being used”

Table 4: Student Comments on Learning the Meaning and Function of Words

The students who reported guessing the meaning of the word from context as their main strategy for learning unknown vocabulary words explained the effectiveness of the DDL activity in some of the same terms as those students who reported using a dictionary as their main strategy for learning new words. For example, these students said DDL helped them to learn the different meanings and forms of the word. However, these students' responses also differed in some respects from those students who use a

dictionary to learn unknown words. While these students find DDL to be similar to their normal strategy of guessing the meaning of new words from context, they reported that DDL is more helpful or effective because: 1) it forces them to think about or guess the meaning of the word for a longer period of time than they would usually devote to such a strategy; 2) it helps them to create sentences using the various forms of the word they find in the language data; and 3) it helps them to notice the range of ways the word can be used, where their usual strategy only allows them to focus on one form and one meaning of the word at a time. These students' responses can be found in Table 5 below.

Miharu	“oh::: it’s more effective + I uhm + I mean I usually look up at the word and guess and so right after that I check the dictionary so I’m not trying to guess the meaning + much longer”
Amir	“I think it’s + they [DDL and guessing the meaning of a word from context] are almost the same [...] [DDL] it goes more than what I usually do so that’s more helpful because usually I just guess the meaning and that’s it but now I learn like as I said like if it’s verb or if it’s noun or if it’s adjective [...] and like uhm take uhm like many + I mean create many sentence about that”
Kamal	“when we read something we focus on one word and we can memorize it in our mind but at the same time is really difficult to like recognize or realize the + the difference between meanings because this maybe this word has different meanings in different situations”
Fardowsa	“[I learned from DDL] that I have confidence that it’s right and that I didn’t need to look up in a dictionary”

Table 5: Student Comments on DDL vs. Guessing the Meaning of Words from Context

Though many students felt they learned something from DDL, not all of the students thought it was helpful. Some students felt that DDL was too difficult, either because the surrounding context contained too many unknown words, or they were unable to guess the meaning for unstated reasons. Furthermore, many students mentioned that they did not like the uncertainty of not knowing the exact meaning of the word, and

those students who did not like to guess the meaning of words from the surrounding context had a particular distaste for this activity. The comments of these students can be found in Table 6.

Basir	“I think the structure for the sentence uhm maybe a little bit difficult to guess the meaning so if you like don’t understand the sentence well so you can’t guess the meaning”
Fardowsa	“uhm learning new words from the dictionary I think it’s better than this because at time when I guess meaning from the reading it may be right or wrong but when I look up in the dictionary the dictionary will let me know maybe my guess is right or wrong”
Nasim	“in this activity uhm and sometimes I think you’re not sure what uhm what does the word means [...] somebody that learning the word in his first time will not be sure what is the exact meaning of the word”
Eri	“I don’t like to guess the meaning from texts like from sentence because I cannot guess the meaning so for me if I find something which I don’t know the words I want to check dictionary as soon as possible so then maybe I can memorize the meaning so yeah I don’t like this one”

Table 6: Student Comments on the Drawbacks of DDL

3.2.3 Results for research question three

RQ3: What are the students’ perceptions of what they are being taught or learning by completing the DDL activity?

Question 7 of the post-activity interview (see Appendix H) asks: “Do you think this activity was designed to teach you something specific about the word *subsequent*? (If yes, what?)”. The responses of those students who felt the activity was designed to teach them something specific about the word can be found in Table 7.

Nasim	“like your trying to teach what is the meaning of the word [...] that was clear I think”
Eri	“uhm there are different kinds of meaning in one word”
Sultan	“uh:m + to know how to guess the word or uhm figure out what’s the uhm what the word means”
Nang	“uhm I think it’s teaching me the uhm the meaning of the word and the part of speech”
Talal	“how to uhm to know uhm part of uh: speech and the meaning of the word from the context”
Maram	“teach them + how the word can be uhm how one word can have many different kind of meaning [...] and also how it can change from verb to adjective”
Hanok	“yeah I feel that the most like common or like useful word that you will use [is what the activity is designed to teach]”

Table 7: Student Perceptions of what the DDL Activity was Designed to Teach

The above data can be placed into five categories of what the students felt they were learning or being taught by completing the DDL activity. The students felt they were learning or being taught: 1) the meaning(s) of the word, 2) the parts of speech of the word, 3) the connection between word form and meaning, 4) the skill of how to guess the meaning from context, and 5) the most common and/or useful words.

3.2.4 Results for research question four

RQ4: In which design do more students notice the collocate *to* of the target word *subsequent*?: When concordance lines are ordered according to the collocate of the target word (Group 1) or when they are ordered at random (Group 2)?

Question 5 of the experiment activity (see Appendices F or G), asked the students to locate any collocates of the word *subsequent* in the data. The activity was designed so that the collocate *to* was a recurring word in the data occurring five times after the word *subsequent*. In the activity given to Group 1, the concordance lines were sorted so that all

the lines with the collocate *to* appeared consecutively, while the concordance lines in the activity given to Group 2 were sorted at random. Neither group outperformed the other on this task. Two students in Group 1 located the collocate while the same number found it in Group 2.

CHAPTER FOUR

DISCUSSION AND CONCLUSION

The following section discusses the results of the study with a focus on the implications for changing and improving the design of the DDL activity used in this study and the unique contributions DDL has to offer to vocabulary teaching in the second language classroom. Before discussing the results, however, I outline where DDL activities of the type featured in this study have their place in the big picture of language instruction as described by I.S.P. Nation (2001). Following the discussion of the results, I end by considering the implications this study has for training learners to work with concordance output.

4.1 Discussion

Before discussing the results of the study and the implications they have for designing and exploiting DDL activities to teach unknown vocabulary, it is important to place the DDL activity featured in this study in its larger context in the communicative language classroom. In the introduction to his book, *Learning Vocabulary in another Language*, I.S.P. Nation (2001) divides the communicative language course into what he calls “four strands” of instruction: 1) comprehensible meaning-focused input, 2) language or form-focused instruction, 3) meaning-focused output, and 4) fluency development. These four strands serve as a guide for balancing instruction in the foreign or second language classroom. Furthermore, Nation claims these strands need to be equally employed in the classroom and in homework in order to create a balance in the language course (i.e., 25% of class/homework time should be spent on each strand). The DDL activity featured in this study clearly falls under strand two of Nation’s model, form-

focused instruction. This said, DDL activities which focus on providing learners with declarative knowledge of the meaning(s), collocate(s), etc. of a word, as did the activity in this study, are merely one part of what should be a more holistic approach to vocabulary teaching and learning; teachers will need to supplement these activities with materials and activities which give learners the opportunity to apply the declarative knowledge of a word they acquire from DDL in meaning-focused and fluency-building activities. One strength of the activity design in this study, however, was that it contained concordance lines which were extracted from the students' assigned texts. Therefore, the students encountered these words in both their meaningful context and in a form-focused, DDL activity.

Keeping in mind that DDL should be considered just one part of a holistic plan to teach academic vocabulary in the EAP classroom (especially once it has unequivocal empirical support for its implementation), the results of this study have many implications for designing and using this type of activity with college-level students. First, the results have implications for adjusting the design of the DDL activity in order to better scaffold learners and promote more accurate and efficient analyses than those performed in this study. Second, the results have implications for the unique applications of DDL in the teaching of unfamiliar vocabulary; specifically, that the activity may facilitate students' "noticing" of certain aspects of the word that traditional vocabulary learning strategies (i.e., dictionary consultation and guessing meaning from context) either do not cover or that students are not able and/or inclined to take advantage of when using these resources. Finally, the results suggest that DDL may also be a viable

technique to teach and test language learners' skill in guessing the meaning of unknown words from context. These results will be discussed further below.

Other studies of DDL activities have found that students find this type of work complicated or difficult to complete (Hadley, 2002). The results of this study corroborate these findings to a certain degree (i.e., no students found the overall activity easy to complete). However, this study offered a more complex view of what students have trouble with when engaging in DDL activities to learn unknown vocabulary. The students in this study did not find the entire DDL activity "difficult" to complete; the majority, 67%, reported that the activity was only moderately hard to complete. What appeared to have influenced the students' perceptions of the difficulty level of this activity were their attitudes toward two of the tasks, which they reported as being the most difficult to complete: 1) identifying the meaning(s) of the word, and 2) writing a definition for the word. These findings suggest that these tasks should be made easier for the students to complete to ensure that they are learning as much about and arrive at as clear of an understanding of the word as possible, and are not spending more time than necessary on these tasks.

One benefit of DDL touted by some applied linguists (Bernardini, 2002; Gavioli, 1997) is that it allows learners to arrive at developmentally appropriate conclusions about the linguistic structures under analysis rather than being forced to provide correct answers to pre-identified language patterns or prefabricated rules. Although I believe that this is indeed a benefit of DDL, there is room for interpretation of this claim. First, in the case of using DDL to teach unknown words, I do not believe that a students' inability to identify the meaning(s) of a word should be classified as developmentally appropriate learning;

this would be too broad of an interpretation. If DDL is going to be used in the classroom to teach unknown vocabulary, teachers need to identify when incomplete descriptions of a word are due to the student's developmental ability versus an error in activity design or inappropriate amount of scaffolding. In this study, only two students were able to approximate the meaning of the word *subsequent*. This is clearly unacceptable. In the case of learning new vocabulary, I suggest that a developmentally appropriate measure of learning should be identified as the way the student explains the meaning, and that this explanation should reflect the meaning(s) of the word. In other words, students' answers should be acceptable in their various forms, as long as their responses arrive at a close approximation of the meaning(s) of the word under analysis.

However, because the design of the DDL activity in this study did not facilitate the majority of the students analyses (i.e., lead them to approximate the meaning(s) of the word *subsequent*) adjustments need to be made in order to ensure that students are able to arrive at a developmentally appropriate understanding of new words from this activity. There are two possible areas where adjustments can be made: training of the students and activity design. Presently, I would like to discuss the latter area. I will address the issue of training students to work with concordance output later in the chapter, as it has implications for more than facilitating students' analyses of concordance lines.

In order to ensure that students arrive at developmentally appropriate answers, the current activity could be amended by supplying the students with definitions (taken from a dictionary appropriate for the learners' proficiency level) of the word and any formulaic language with unique meanings which appear in the data. The students can then categorize the concordance output according to these definitions. However, so that

inductive learning and guessing the meaning from context is still encouraged, and all the benefits of these practices are maintained, teachers can wait to supply their learners with definitions until a second round of analysis. This way, students should undoubtedly arrive at a clear and approximate understanding of the meaning and use of the new word while still being able to practice and possibly improve their skill in guessing the meaning of a word from its context, and reap the benefits of the more traditional DDL activity design, which has been found to be interesting (Hadley, 2002), fun, and useful (Götz & Mukherjee, 2006).

This change in design should not be implemented without question, however. Further research is needed to determine how this change alters the DDL experience for students, and the implications it has for language acquisition. However, Chambers (2005) has found that learners “can derive benefit from corpus consultation without realizing the full potential of the activity” (p. 117). In other words, learning has been documented by researchers who combine inductive and deductive analyses in DDL. Furthermore, altering the activity in this way will undoubtedly lead to the students spending less time on this portion of the activity, which will allot more time for other tasks (only six students or 40% were able to complete the entire activity in forty-five minutes). Finally, scaffolding students’ analyses with definitions from the dictionary should help them to have confidence in their answers. During the interview many students expressed that they were unsure of the correctness of their answers (see Table 6). Some of this anxiety about correctness may be unavoidable in DDL because its design implicitly creates a new learning environment for students where correctness is no longer narrowly defined, and students are forced to deal with language in all its natural complexity – an understandably

daunting task. However, providing definitions of a word and its formulae can only help to mitigate students' anxiety, frustration, and feelings that the DDL activity is difficult to complete, which should have the effect of increasing motivation.

The results of the students' comparison of DDL with their usual strategies for learning new words (i.e., research question two) indicate that the students perceive that they are learning something from DDL that they would not have if they had used their traditional methods for learning new words. Furthermore, what they reported learning may have important implications for the way language is taught and acquired.

Only two studies, Boulton (2008) and Chambers (2005) have explicitly focused on how concordance output can best be integrated into the language classroom by having students compare corpus-based activities to traditional methods for learning vocabulary and grammar, such as by consulting a grammar, dictionary, or course book. The results of the current study will be discussed in light of these two studies, where relevant, below.

Chambers' (2005) study asked undergraduate students to compare the usefulness of direct corpus consultation with that of a course book, dictionary or grammar while Boulton (2008) compared the comments of his second-year, architecture students, one group using corpus data and DDL and a second group using dictionary entries and traditional teaching methods to learn problematic grammar items as identified in the students' own writing prior to the experiment. The current study differs from both these experiments in that it asked students to compare DDL with their own methods for learning new words rather than traditional reference materials or teacher initiated, classroom techniques. This led to the students comparing DDL to skills-based techniques, namely guessing the meaning of words from context, as well as traditional

reference materials, which in this study was limited to print and electronic dictionaries. Furthermore, the current study was narrowly focused on the students' opinions of what they felt they learned from DDL that they believe they could not have learned from their usual methods of learning new words. The students were not asked to give value judgments of the activity, as many studies have previously found that students rate DDL positively (Hadley, 2002; Chambers, 2005; Götz & Mukherjee, 2006; Boulton, 2008).

When students in the current study compared DDL to their usual methods for learning new words, they remarked that DDL helped them to notice certain things about the word that they would not have using their traditional methods. When compared to dictionary work, students commented that DDL caused them to realize the connection between the form of the word and its meaning; for example, one student, Hanok, commented:

in the dictionary it's crazy they give you like *subsequent* and they give you the adverb the adjective and this and that but these things are used in various sentences so that you can see okay in this sentence it's used as a noun in this sentence it is used as an adverb in the sentences it is used as an adjective in this sentence it is used as a verb + now you'll be able to figure it out okay right now I can tell okay this is a verb there is the way it can be used [...] there is different from that of the dictionary

Although these students did not make any explicit remarks about how certain forms of the word *subsequent* determined its meaning, their comments suggest that these students have realized that form and meaning are connected. This finding is quite surprising as the researcher/instructor neither lectured nor commented on this relatively new conception of language most espoused by John Sinclair (1991) after his involvement in the Cobuild project:

The recognition that form is often in alignment with meaning was an important step [...] Soon it was realized that form could actually be a

determiner of meaning, and a causal connection was postulated [...] Then a conceptual adjustment was made, with the realization that the choice of meaning, anywhere in a text, must have a profound effect on the surrounding choices. (Sinclair, 1991, p. 7)

Though it cannot be strongly confirmed based on the data collected in this study, it is quite possible that the design of the activity lead students to this conclusion about the connection between form and meaning, as students were asked to first identify the parts of speech they found in the data and then to use these categories to determine the meaning(s) of the word. Furthermore, 53% (n=8) of the students answered that “grammar and vocabulary are two completely separate parts of the English language” in the Grammar and Vocabulary Questionnaire (see Appendix E) they filled out prior to taking part in the experiment. This percentage includes every student (n=4) who mentioned making this connection in the post-activity interview (see all of Table 3 and Maram’s comments in Table 7). These data are further corroborated by Chambers (2005), who found that when consulting a corpus, learners “often made discoveries which were lexico-grammatical rather than solely grammatical in nature” (p. 116) despite beginning their analyses with traditional, grammatical categories. In light of this evidence, appears that DDL may play a role in promoting language learners’ noticing of this new and (as more research continues to shed light on the form-meaning relationship of words) undoubtedly important, linguistic phenomenon. However, this study only identified that students noticed this connection; it did not examine students’ knowledge of the form-meaning relationship of the word *subsequent*. Future research needs to examine how students actually acquire the language once they understand this connection between form and meaning, and how this type of acquisition differs from students who have not

realized this relationship and/or have only utilized traditional methods of learning words (e.g., with wordlists and dictionaries).

In addition to noticing that there is a relationship between the form of a word and its meaning, students who normally used a dictionary to learn new words commented that DDL helped them to notice the different forms (i.e., inflections and derivations) of the word and the different senses of the word. They commented that learning many forms and senses of the word at one time was something they did not do when consulting a dictionary. Instead, they use a dictionary to obtain information about the specific form or meaning of the word that is relevant to them at the time. For example, one student, Maram, commented on learning the different forms while, Amir, shares his thoughts on learning the word's meanings:

Maram: it [DDL] teach me also uhm uhm (1) if I check a word in dictionary I just understand that it is verb I don't understand that for I don't check up or down of that uhm word that understand that what is the adjective of that word or noun or adverb but here I can understand that it has how many + how many branches that one word has

Amir: some words they have like many or a lot of meanings or like two or three meanings + so that usually I just take one meaning and that's it but with I mean by this way + I learn like three or four meanings

Interestingly, these students' do not fault reference materials for providing incomplete information about the language as the learners do in Chambers (2005) study. Instead, the students' characterize the benefits of DDL in contrast with their own habits of using reference materials. In other words, it is not that the students believe that a dictionary cannot provide them with the information that DDL does, at least they never explicitly state such a belief, but that the DDL activity pushes them to learn more information about a word than they are motivated or likely to when studying vocabulary

on their own. This begs the question: Does learning multiple meanings and all the forms of an English word at one time better facilitate acquisition of the language than more traditional methods of learning new words? Future research might compare student's productive use of words when learned through DDL vs. traditional methods in order to establish if DDL leads to a greater recall, range of use and/or native-like command of the word.

Finally, students felt DDL was more useful than a dictionary because it helped them to understand how words are used to create meaning (i.e., their function). These students condemned the isolated way words are displayed in traditional dictionaries and were enthusiastic about being exposed to the word in a variety of contexts in DDL; for example, Hanok commented:

the dictionary will just + just drop the word BOMB and then give the definition but this + look at this activity you have you have fifteen different ways in which the word is used [...] now it give you broader idea of what the meaning or what the word should be [...] and how I should be able to use it in a sentence [...] I think the bottom line is uhm I realize that words can be different the dictionary will give the word but that particular word somebody can use in different way [...] you will not just see the word and just say there is the exact meaning unless it is being used

Hanok's comment is a nice paraphrase of one made by John Sinclair (1991) in his book *Corpus Concordance Collocation*:

The models of meaning that we are 'given' by linguistic tradition are the dictionary and the thesaurus. The traditional dictionary cheerfully represents words as often having several discrete meanings, but gives no help whatever as to how in practice the language user distinguishes among them – how a writer can be fairly sure that the meaning he or she wants to signal is the one which will be understood, and vice versa. (p. 7)

Boulton (2008) also found that his learners felt corpora were “most useful for the contexts and ‘concrete examples’ which highlight usage” (p. 42).

Although the context in which words appear in concordance output is limited (often to seven or less words on either side of the target word), and this often occludes information about how the word figures into the overall discourse structure of a text, students in and beyond this study still report that they are able to extract meaning from the concordance lines; meaning which is valuable from a bottom-up perspective of discourse rather than top-down. In this study, some students felt that they acquired an understanding of how the word is used, “I realize that words can be different the dictionary will give the word but that particular word somebody can use in different way”, and that this would help them to use the word in the future, “I should be able to use it in a sentence”. To a certain extent, dictionaries are beginning to feature more authentic examples of words in context, but as other studies have shown (Chambers, 2005; Walker, 2009), the information is often incomplete. Furthermore, the results of this study indicate that learners often do not use dictionaries to teach themselves multiple aspects of a word. Therefore, DDL appears to be the most advantageous option for exposing students to a word in multiple, authentic contexts at one time. However, again, future research will need to verify if this is indeed a favorable condition for language acquisition, and if the meaning students report finding in the data actually helps them to use the word when composing longer texts within a given discourse.

In addition to delineating clear advantages of using DDL over consulting a dictionary to learn new words, most learners in this study commented that DDL was a more effective activity for learning new words than guessing the meaning of a word from a single context. These students felt that DDL was similar to guessing the meaning from a single context, but that it was a more effective strategy because it prompted them to guess

the meaning of the word for a longer period of time, it gave them access to multiple meanings of a word at one time instead of a single meaning, and one student remarked that it gave her greater confidence in her ability to guess the meaning of a word from the context. These findings are unique in that no other study reviewed has sought to delineate the uses of DDL over guessing the meaning of words from context. The implications of these findings are that DDL could serve an ancillary purpose to teaching unknown words and be used to teach, strengthen, and test students' skill in guessing the meaning of unknown words from context; a skill thought to be vital to academic success and which can be difficult for students to perfect. Furthermore, DDL may increase students' confidence in their ability to utilize this vocabulary learning strategy, hopefully leading to more efficient and autonomous use of this strategy outside of the classroom. The only caveat being that students who do not like to guess the meaning of words from context before engaging in DDL may be further frustrated by and unmotivated to engage in the activity, as was discovered through students' comments on the drawbacks of DDL (see Table 7) in this study.

A final area of DDL examined was the effect of two designs of the activity on students' noticing of collocations. To my knowledge, this is the first study to examine design implications for this potential use of DDL, although many researchers using corpus-based activities with learners have touted it for highlighting such linguistic structures (Bernardini, 2002; Gavioli, 1997, 2001; Thurston & Candlin, 1998). The results show that neither design (i.e., concordance lines arranged according to the target word's collocate or randomly) lead to more students noticing the collocate *to* in the data. Furthermore, 87% (N=13) of the students failed to locate the collocate altogether. These

results suggest that there are other important factors which prevented the students from identifying the collocate in the data. Although these students were given a short lecture, accompanied by a hand-out, on formulaic language (see Appendix C), this information was not sufficient in helping them to identify collocates. Furthermore, no student mentioned that they thought the activity was designed to teach them the collocations of the target word in the post-activity interview (see the results to research question three or Table 7 above). This was an unexpected result of the study because it was assumed that the short lecture on the topic (see Appendix C), the inclusion of task number five in the activity (see Appendix D), and the activity's design would facilitate noticing of this linguistic structure. Future research is needed to investigate why students do not notice collocations in DDL activities if students are to fully benefit from these activities. In turn, learning to identify collocations in concordance data may be a transferable skill, allowing them to more quickly identify such sequences on their own (i.e., outside of the classroom) and thus, acquire a more native-like competency of the language.

One aspect of DDL not explicitly examined that had a wide effect on the design and results of the study was training the students to perform analyses on the concordance output. I am not the first to discuss issues of training students to categorize the language structures they find in concordance data. However, the results of this study confirm, once again, that training students to analyze and categorize concordance lines is a time-consuming and, in the case of teaching them to look for formulaic language, difficult process. The students involved in this study were given lectures on the basics of corpus analysis, formulaic language, and given the opportunity to analyze concordance output during three, 45-minute sessions. All together, training the students amounted to 3^{1/2} hours

of instruction. This is not an exorbitant amount of time, but when one considers that these students still found it difficult to categorize the data according to the word's meanings and to identify collocates after this training, it is clear that the issue of training needs to be resolved before teachers should be encouraged to use DDL in their classrooms, despite all the benefits it has found to offer in this and other studies. Researchers need to determine, through empirical studies, if it is appropriate for EAP, or indeed other, instructors to block off valuable class time to train students to work with concordance output. In other words, what is the cost-benefit or return for learning of students who engage in DDL to learn new words? If, through empirical studies, DDL is proven to be a technique which is more efficacious than other techniques in promoting language acquisition and able to teach aspects of the language that other resources or techniques are unable to, and indeed I believe it will, then the question of when and where it is appropriate to train learners will become extremely important. Given the fact that it is a time-consuming process, and that the students in this study reported that they did not get the full benefit of language reference materials and had trouble with guessing the meaning of words from context, it may be appropriate to offer a one-credit course to ESL students, early in their college career, where they would receive training in vital language learning skills, such as guessing meaning from context, utilizing language reference materials, and using a corpus. This is assuming that language instructors are facing a similar situation with their ESL students in other institutions of higher learning in the U.S. If such a course were offered, this would also mean that principles of corpus design would have to be reconsidered so that freely available pedagogical corpora could

be accessed by these institutions and not build from scratch to meet these students' purposes; corpora that would perhaps divide its texts based on topic and reading level.

4.1 Conclusion

The main goal of this exploratory study was to examine if DDL has something unique to offer in the teaching of vocabulary, warranting its use alongside more traditional vocabulary teaching methods. Furthermore, the study compared the efficacy of two designs on students' noticing of collocations in concordance data. Finally, suggestions were made for improving the overall design of the activity in order to better scaffold students' analyses and make DDL work easier for students to complete.

Appendix A: Student Background Information

Student	Group #	Age	Class Standing	Major(s)	Country of Origin	Native Lg(s).	Second Lg(s).	Years studying English	Years in the U.S.
Zia	1	22	Freshman	Respiratory Therapy	Saudi Arabia	Arabic	English	2	3
Maram	1	22	Junior	Computer Engineering	Iran	Farsi	English	2	0
Kwan	2	24	Freshman	Sports Management	South Korea	Korean	?	9/12	1
Basir	2	20	Freshman	Computer Science	Saudi Arabia	Arabic	English	2	2
Nasim	2	23	Sophomore	Mechanical Engineering	Pakistan	Arabic	English	approx. 1 year	2 weeks
Talal	1	28	Sophomore	Biology	Saudi Arabia	Arabic	English	3	1
Waleed	2	22	Sophomore	?	Saudi Arabia	Arabic	English	2	2
Eri	1	20	Sophomore	Business	Japan	Japanese	English	8	5/12
Sultan	1	22	Freshman	Finance	Kuwait	Arabic	English	3	1.5
Fardowsa	2	25	Freshman	Computer Info. Tech.	Nigeria	Yoruba	English	10	1.3
Nang	1	25	Freshman	Electrical Engineering	Saudi Arabia	Arabic	English	8	1.3
Kamal	1	20	Freshman	Electrical Engineering	Saudi Arabia	Arabic	English	1 5/12	1.5
Amir	2	21	Freshman	Finance	Saudi Arabia	Arabic	English	1	1
Hanok	2	?	Sophomore	International Studies	Liberia	Kru	English	?	1
Ayano	1	20	Sophomore	Business	Japan	Japanese	English	8	5/12

? = Unknown information

Appendix B: List of AWL Headwords for G011 Course

acquire	distinct	legal
adequate	distribute	locate
adjust	economy	major
alternative	emerge	method
analyze	enable	notion
approach	environment	occur
area	establish	percent
assess	estimate	period
assume	evaluate	policy
authority	evident	primary
available	facilitate	principle
benefit	factor	proceed
civil	finance	range
component	formula	require
concept	identify	section
consist	image	sector
context	implement	seek
create	indicate	significant
data	individual	similar
debate	initial	source
decline	interpret	specific
deduce	involve	structure
define	issue	sufficient
derive	labor	theory
		vary

Appendix C: Introduction to Formulaic Language Hand-Out

Introduction to Formulaic Language: Free Combinations, Collocations, and Idioms

Overview of Formulaic Language

It is estimated that 80% of the English language could be formulaic (Wray, 2000); therefore, it is important to familiarize yourself with certain types of formulaic language if:

- You are concerned about being quickly and clearly understood by native speakers of English and vice versa
- You want to improve your reading speed and comprehension in English
- You want to gain more clarity in your writing in English

Formulaic language: Language which is stored in the mind and retrieved or recalled whole when you are using the language to write or speak as opposed to using words and grammar rules to form sentences and phrases. These formulaic phrases are also sometimes called “chunks” of language.

You probably already know many forms of formulaic language without knowing it; for example, the phrase *How are you?* is a formulaic sequence. However, formulaic language is more complicated than just stock phrases or clichés.

Three Levels of Language:

In particular, it is important to distinguish between three levels of formulaic language, which range from word combinations that are not formulaic to combinations that are completely fixed. The three levels are: free combinations, collocations, and idioms.

Free Combinations: A free combination occurs when there is no restriction on any of the words in the phrase. These phrases can be created by knowing the meaning of individual words and applying grammar rules to put them together (see the example below).

Example of a free combination: The girl saw the boy.

This is an example of a free combination because there are many words that could replace the noun, *girl*, the verb, *saw*, and the noun *boy*. Furthermore, this sentence was created by knowing the meaning of *girl*, *saw*, and *boy* and using basic grammar knowledge to put the words in the correct place in the sentence.

Collocations: collocations are words that frequently occur together or have a greater than chance probability of occurring together, and seem to be “glued together” in the mind of the native speaker. For example, the phrase *crystal clear* is a common adjective-noun collocation in the English language. The main characteristic of collocations that make them different from idioms is that the meaning of the entire phrase can be derived from the meaning of the individual words in the phrase. It is important that you know two types of collocations: *grammatical collocations* and *lexical collocations*.

Grammatical Collocations: collocations which consist of a noun, an adjective, or a verb, plus a preposition or a grammatical structure such as an infinitive or clause (Bahns, 1993).

Type of collocation	Examples	Parts of speech
Phrasal verbs	consist of account for	Verb + Preposition (particle)
No name	afraid/scared of strong in subject to	Adjective + Preposition
No name	strength to lift continued to work struggled to find	Noun + (a particular form of the)Verb

Lexical Collocations: collocations which consist of various combinations of nouns, adjectives, verbs, and adverbs (Bahns, 1993).

Examples	Parts of Speech
completely satisfied absolutely gorgeous utterly ridiculous	Adverb + Adjective
excruciating pain rancid butter big mouth	Adjective + Noun
lion's roar	Noun + Verb
commit suicide/murder consult a doctor	Verb + Noun

Idioms: An idiom is a phrase that is characteristic of a particular language. Furthermore, an idiom cannot necessarily be fully understood from the separate meanings of the individual words which form it as collocations can, but instead must be learned as a whole unit of meaning. Idioms are sometimes difficult to understand because they sometimes rely heavily on an understanding of the culture of the target language for their meaning. There are more transparent (i.e. clear) idioms, such as, *looking for a needle in a haystack*, where the meaning of the phrase can be understood by the meaning of the words in it (here *looking for a needle in a haystack* means to look for an item that is very difficult or impossible to find). However, there are also more opaque (i.e. not clear) idioms, such as, *to have the upper hand*, where the words in the phrase do not reveal the meaning of the phrase (here *to have the upper hand* has nothing to do with "hands", but means to have a position of control or advantage over someone or something).

Type of Idiom	Examples
Transparent Idioms	<i>in the know</i> : to have knowledge of something, especially secret or special information <i>(to) get your money's worth</i> : to get good value for the money you spend on a product
Opaque Idioms	<i>(to) beat around the bush</i> : to be deliberately ambiguous or unclear in order to mislead someone or to withhold information <i>(to) bring home the bacon</i> : to earn money to live on

Appendix D: DDL Training Activity

*1 retirees to workers. As European population growth **declines**, and as immigration increases, the
*2 ort from the East/West Center in Hawaii notes the **decline** in family and authority in Asia, and
3 less than eighteen and a half hours per week. The **decline** in hours spent studying has to be
4 groups. The percentage of females who drink in cars **declines** rapidly with age (from 40.6 to 27.6
5 not otherwise have been possible. Although Britain **declined** as a great power during the 1960s
6 in our universities, the institution itself can rapidly **decline**, as it did in the seventeenth and eig
7 At the end of the year the number of detainees had **declined** to about 60. 2 Two factors make
8 verage decline of 3.5 Tcf. Natural gas consumption **declines**, in both the lower oil price and low
9 unrest at that time. The Anarchist movement was **declining**, and gradually being replaced by
10 While the number of work permits was **declining** by the early 1970s due to the

* Sentences taken from *Caught Between the Ages* in your Sociology textbook.

- 1) Categorize or separate the word according to its part of speech (noun, adjective, verb, adverb, prepositions, verbals, etc...).

- 2) Categorize or separate each part of speech according to its meaning.

- 3) How many meanings for the word did you find?

- 4) Give a definition for each meaning you found in the data. Go back to question two in order to make a list of definitions.

- 5) Look at the language data again according to its part of speech. This time locate any recurring patterns in the data that you think are idiomatic phrases or collocations.

- 6) Create one original sentence for each meaning of the word.

Appendix E: Vocabulary and Grammar Questionnaire

Directions: Put a check in the box that most accurately matches how you feel about each statement. Please check only one box for each statement. For question 14, please finish the statement by telling me how you learn or study unknown English words.

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
1. I like to study English vocabulary					
2. The best way for me to learn new words is to look them up in a dictionary					
3. The best way for me to learn new words is to read a lot					
4. The best way for me to learn new words is to guess their meaning from the context					
5. I like to study English grammar					
6. The best way for me to learn English grammar is to study English grammar rules					
7. The best way for me to learn English grammar is to read a lot in English					
8. The best way for me to learn English grammar is to study English grammar rules and read a lot in English					
9. Grammar and vocabulary are two completely separate parts of the English language					
10. Grammar and vocabulary should be taught in separate lessons					
11. I must know grammar rules in order to speak or write in a language					
12. I must know many individual words in order to speak or write in a language					
13. I must know many word chunks, like <i>drop off</i> , in order to speak or write in a language					
14. When I don't know a word I:					

Appendix F: Experiment Session DDL Activity (Group 1)

1 itself has adverse consequences for subsequent opportunities. In particular, incarceration is
2 selective effects, severely limits subsequent employment opportunities. And while the audit
3 the "service of humanity." On subsequent visits to England, Miss Adams was to meet many of these people, but
4 graduates, both in first year classes and subsequent college and university graduating classes" (Joint Committee for Review of
5 and set it on fire. It does not matter that subsequent studies of the riot showed that explanation was nonsense. 4 Members of the
6 free-throws in large groups, free-throws subsequent to the first two attempts (i.e. the majority of practice attempts) are quite
7 Such research could be completed subsequent to a similar event by assessing level of identification with the athletes in
8 an added benefit the debate that arose subsequent to his decision helped the campus community to discern the error in one
9 an academic term with little follow-up subsequent to teaching episodes. Using end-of-the semester student course evaluations
10 students' literacy skills prior to and subsequent to admission and if skills are weak, faculty should be concerned about
11 give information about a topic and then subsequently state their opinions about how the topics affect them, which is a value
12 explained to the student teacher at a subsequently scheduled conference time. As the student teacher begins to plan
13 staff or by a single music teacher who subsequently obtains the approval of the administration. Independent study, whether
14 one week later. However, she was subsequently admitted to another hospital with continuing vaginal bleeding. An
15 for the past three years and have subsequently identified elements that lead to a successful team-teaching experience

- 1) Categorize or separate the word according to its part of speech (noun, adjective, verb, adverb, prepositions, verbals, etc...).
- 2) Categorize or separate each part of speech according to its meaning.
- 3) How many meanings for the word did you find?
- 4) Give a definition for each meaning you found in the data. Go back to question two in order to make a list of definitions.
- 5) Look at the language data again according to its part of speech. This time locate any recurring patterns in the data that you think are idiomatic phrases or collocations.
- 6) Create one original sentence for each meaning of the word.

Appendix G: Experiment Session DDL Activity (Group 2)

1 an added benefit the debate that arose subsequent to his decision helped the campus community to discern the error in one
2 and set it on fire. It does not matter that subsequent studies of the riot showed that explanation was nonsense. 4 Members of the
3 explained to the student teacher at a subsequently scheduled conference time. As the student teacher begins to plan
4 for the past three years and have subsequently identified elements that lead to a successful team-teaching experience
5 free-throws in large groups, free-throws subsequent to the first two attempts (i.e. the majority of practice attempts) are quite
6 give information about a topic and then subsequently state their opinions about how the topics affect them, which is a value
7 graduates, both in first year classes and subsequent college and university graduating classes" (Joint Committee for Review of
8 itself has adverse consequences for subsequent opportunities. In particular, incarceration is
9 Such research could be completed subsequent to a similar event by assessing level of identification with the athletes in
10 one week later. However, she was subsequently admitted to another hospital with continuing vaginal bleeding. An
11 s elective effects, severely limits subsequent employment opportunities. And while the audit
12 staff or by a single music teacher who subsequently obtains the approval of the administration. Independent study, whether
13 students' literacy skills prior to and subsequent to admission and if skills are weak, faculty should be concerned about
14 the "service of humanity." On subsequent visits to England, Miss Adams was to meet many of these people, but
15 an academic term with little follow-up subsequent to teaching episodes. Using end-of-the semester student course evaluations

- 1) Categorize or separate the word according to its part of speech (noun, adjective, verb, adverb, prepositions, verbals, etc...).
- 2) Categorize or separate each part of speech according to its meaning.
- 3) How many meanings for the word did you find?
- 4) Give a definition for each meaning you found in the data. Go back to question two in order to make a list of definitions.
- 5) Look at the language data again according to its part of speech. This time locate any recurring patterns in the data that you think are idiomatic phrases or collocations.
- 6) Create one original sentence for each meaning of the word.

Appendix H: Post-Activity Interview Protocol

Name of Interviewee: _____ Time: _____

Part 1: Perceptions of the Difficulty of the Activity

Ask student's to think only about the activity they just completed (i.e., they should disregard the training session activities) to answer the questions in this section. Then, say: "This first set of questions is meant to measure your perceptions of the difficulty of the different tasks of this activity. For each question, your answer choices are: easy, moderate, or difficult (explain the word moderate if necessary). Please answer each question as honestly and accurately as possible. There are no right or wrong answers. Please select only one response for each question.

1. How would you rate the overall difficulty of this entire activity?

Easy	Moderate	Difficult
------	----------	-----------

2. How difficult was it for you to identify the part(s) of speech of the vocabulary word in the data?

Easy	Moderate	Difficult
------	----------	-----------

3. How difficult was it for you to identify the meaning(s) of the vocabulary word in the data?

Easy	Moderate	Difficult
------	----------	-----------

4. How difficult was it for you to write a statement (definition) about the meaning of the new word?

Easy	Moderate	Difficult
------	----------	-----------

5. How difficult was it to create an original sentence using the new word?

Easy	Moderate	Difficult
------	----------	-----------

6. How do you feel about learning new words in this way?

Part 2: Student Perceptions of What They Learned

In the next set of questions, I am interested in hearing about how you feel about the structure of this activity, including what you feel you have learned by completing this activity and how you searched the data. Please be as honest as you can in your responses. There are, of course, no right or wrong answers.

7. Do you think this activity was designed to teach you something specific about the word *subsequent*? (If yes, what?)

Part 3: Comparison of Activity with Traditional Vocabulary Acquisition Strategies

In this last set of questions, I am interested in your perceptions of the difference between your traditional method(s) for learning new vocabulary and the activity you just completed. Remind the interviewee of his/her strategy for learning new words as outlined in their response to the question “when I don’t know a word I _____” from Grammar and Vocabulary questionnaire.

8. How do you think today’s activity compares to your traditional method of learning new vocabulary words?

Prompts, if needed:

- Do you think you learned more or less about the word? Explain.
- Is it more or less difficult to learn new words this way? Explain.
- What do you think are the drawbacks/benefits of learning new words this way?

9. What did you learn, if anything, about the vocabulary word that you might not have if you taught yourself the word using your normal strategy for learning new words?

REFERENCES

- Anthony, L. (2008). *AntConc (Version3.2.1.)*[computer software]. Retrieved November 1, 2008, from <http://www.antlab.sci.waseda.ac.jp/software.html>.
- Aston, G. (1995). Corpora in language pedagogy: Matching theory and practice. In G. Cook & B. Seidlhofer (Eds.), *Principle & practice in applied linguistics: Studies in honour of H.G. Widdowson*. Oxford, UK: Oxford University Press.
- Aston, G. (1997). Enriching the learning environment: Corpora in ELT. In A. Wichmann, S. Fligelstone, T. McEnery, & G. Knowles (Eds.), *Teaching and language corpora* (pp. 51-64). London: Longman.
- Bahns, J. (1993). Lexical collocations: A contrastive view. *ELT Journal*, 47, 56-63.
- Bernardini, S. (2002). Exploring new directions for discovery learning. In B. Kettemann & G. Marko (Eds.), *Teaching and learning by doing corpus analysis* (pp. 165-182). Amsterdam: Rodopi.
- Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*. Cambridge: Cambridge University Press.
- Boulton, A. (2008). DDL: Reaching the parts other teaching can't reach? In Frankenburg-Garcia, A. (Ed.), *Proceedings of the 8th Teaching and Language Corpora Conference (TaLc 8)* (pp. 28-44). Amsterdam: Rodopi.
- Brown, D.F. (1974). Advanced vocabulary teaching: The problem of collocation. *RELC Journal*, 5, 1-11.
- Cantos, P., & Sánchez, A. (2001). Lexical constellations: What collocates fail to tell. *International Journal of Corpus Linguistics*, 6(2), 199-228.
- Chambers, A. (2005). Integrating corpus consultation in language studies. *Language Learning & Technology*, 9, 111-125.
- Chambers, A. (2007). Popularising corpus consultation by language learners and teachers. In E. Hidalgo, L. Quereda, & J. Santana (Eds.), *Corpora in the FL classroom: Selected papers from the sixth international conference on teaching and language corpora (TaLC6)* (pp. 3-16). Amsterdam: Rodopi.
- Cobb, T. (1997). Is there any measurable learning from hands-on concordancing?. *System*, 25, 301-315.
- Cobb, T. (1999). Breadth and depth of lexical acquisition with hands-on concordancing. *Computer Assisted Language Learning*, 12, 345-360.
- Cobb, T. (2007). Computing the vocabulary demands of L2 reading. *Language Learning & Technology*, 11, 38-63.
- Coniam, D. (1994). Using corpus word frequency data in the automatic generation of English language cloze tests. In L. Flowerdew & A.K.K. Tong (Eds.), *Entering Text* (pp. 29-44). Hong Kong: Language Centre.
- Conrad, S. (2005). Corpus linguistics and L2 teaching. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 393-409). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.
- Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, 34, 213-238.
- Coxhead, A. (2002). The academic word list: A corpus-based word list for academic purposes. In B. Kettemann & G. Marks (Eds.), *Teaching and language corpora (TaLC) conference proceedings* (pp. 252-267). Cambridge: Cambridge University Press.

- Davies, M. (2008-). The Corpus of Contemporary American English (COCA): 400+ million words, 1990-present. Available online at <http://www.americancorpus.org>.
- Ellis, N.C., Simpson-Vlach, R., & Maynard, C. (2008). Formulaic language in native and second language speakers: Psycholinguistics, corpus linguistics, and TESOL. *TESOL Quarterly*, 42, 375-396.
- Fan, M. (2009). An exploratory study of collocational use by ESL students-A task based approach. *System*, 37, 110-123.
- Farghal, M., & Obiedat, H. (1995). Collocations: A neglected variable in EFL. *International Review of Applied Linguistics in Language Teaching*, 33(4), 315-327.
- Flowerdew, J. (1993). Concordancing as a tool in course design. *System*, 21, 231-244.
- Gardner, D., & Davies, M. (2007). Pointing out frequent phrasal verbs: A corpus-based analysis. *TESOL Quarterly*, 41 (2), 339-359.
- Gavioli, L. (1997). Exploring texts through the concordancer: Guiding the learner. In A. Wichmann, S. Fligelstone, T. McEnery, & G. Knowles (Eds.), *Teaching and language corpora* (pp. 83-99). London: Longman.
- Gavioli, L. (2001). The learner as researcher: Introducing corpus concordancing in the classroom. In G. Aston (Ed.), *Learning with corpora* (pp. 108-137). Houston, TX: Athelstan.
- Götz, S., & Mukherjee, J. (2006). Evaluation of data-driven learning in university teaching: A project report. In S. Braun, K. Kohn, & J. Mukherjee (Eds.), *Corpus technology and language pedagogy* (pp. 49-67). Frankfurt, DE: Peter Lang.
- Henslin, J.M. (2007). *Down to earth sociology: Introductory readings (14th ed.)*. New York: Free Press.
- Hadley, G. (2002). An introduction to data-driven learning. *RELC journal*, 33, 99-124.
- Horst, M., Cobb, T., & Nicolae, I. (2005). Expanding academic vocabulary with an interactive on-line database. *Language Learning & Language*, 9, 90-110.
- Johns, T. (2000a). From printout to handout: Grammar and vocabulary teaching in the context of Data-driven learning. In T. Johns & P. King (Eds.), *Classroom concordancing*, *ELR Journal*, 4, 27-46.
- Johns, T. (2000b). Should you be persuaded: Two examples of Data-driven learning. In T. Johns & P. King (Eds.), *Classroom concordancing*, *ELR Journal*, 4, 1-16.
- Johns, T. (2002). Data-driven learning: The perpetual challenge. In B. Kettemann & G. Marko (Eds.), *Teaching and learning by doing corpus analysis* (pp. 107-117). Amsterdam: Rodopi.
- Kennedy, G. (2003). Amplifier collocations in the British National Corpus: Implications for English language teaching. *TESOL Quarterly*, 37, 467-487.
- Kettemann, B. (1995). On the use of concordancing in ELT. *Arbeiten aus Anglistik und Amerikanistik*, 20, 29-41.
- Leech, G. (1997). Teaching and language corpora: A convergence. In A. Wichmann, S. Fligelstone, T. McEnery, & G. Knowles (Eds.), *Teaching and language corpora* (pp. 1-23). London: Longman.
- Lennon, P. (2005). Lexical collocation and the teaching of lexical meaning. *Neusprachliche Mitteilungen Aus Wissenschaft Und Praxis*, 58, 18-24.

- Man-lai, A.C., Pui-yiu, K.W., & Chau-ping, M.W. (1994). Collocational problems amongst ESL learners: A corpus-based study. In L. Flowerdew & A.K.K. Tong (Eds.), *Entering Text* (pp. 157-165). Hong Kong: Language Centre.
- Meyer, C.F. (2002). *English corpus linguistics: An introduction*. Cambridge: Cambridge University Press.
- Meyer, C.F. (2009). In the profession: The “empirical tradition” in linguistics. *Journal of English Linguistics*, 37, 208-213.
- Mukherjee, J. (2006). Corpus linguistics and language pedagogy: The state of the art – and beyond. In S. Braun, K. Kohn, & J. Mukherjee (Eds.), *Corpus technology and language pedagogy* (pp. 5-24). Frankfurt: Peter Lang.
- Nation, I.S.P. (2004). *Learning vocabulary in another language* (5th ed.). Cambridge: Cambridge University Press.
- Pickard, V. (1994). Producing a concordance-based self-access vocabulary package: Some problems and solutions. In L. Flowerdew & A.K.K. Tong (Eds.), *Entering Text* (pp. 215-226). Hong Kong: Language Centre.
- Shillaw, J. (1994). Using a corpus to develop vocabulary tests. In L. Flowerdew & A.K.K. Tong (Eds.), *Entering Text* (pp. 166-182). Hong Kong: Language Centre.
- Simpson, R. & Mendis, D. (2003). A corpus based study of idioms in academic speech. *TESOL Quarterly*, 37(3), 419-441.
- Sinclair, J. (1991). *Corpus Concordance Collocation*. Oxford: Oxford University Press.
- Sinclair, J. (2003). *Reading concordances*. London: Pearson Education Limited.
- Siyanova, A. & Schmitt, N. (2008). L2 learner production and processing of collocation: A multi-study perspective. *The Canadian Modern Language Review*, 64(3), 429-458.
- Sripicharn, P. (2004). Examining native speakers’ and learners’ investigation of the same concordance data and its implications for classroom concordancing with ELF learners. In G. Aston, S. Bernardini, & D. Stewart (Eds.), *Corpora and language learners* (pp. 233-245). Amsterdam: John Benjamins Publishing Company.
- Stevens, V. (1991). Classroom concordancing: Vocabulary materials derived from relevant, authentic text. *English for Specific Purposes*, 10, 35-46.
- Stevens, V. (2000). Concordance-based vocabulary exercises: A viable alternative to gap-fillers. In T. Johns & P. King (Eds.), *Classroom concordancing*, *ELR Journal*, 4, 47-62.
- Sun, Y. & Wang, L. (2003). Concordancers in the EFL classroom: Cognitive approaches and collocation difficulty. *Computer Assisted Language Learning*, 16(1), 83-94.
- Thurstun, J., & Candlin, C. N. (1998). Concordancing and the teaching of the vocabulary of academic English. *English for Specific Purposes*, 17(3), 267-280.
- Tribble, C. & Jones, G. (1990). *Concordances in the classroom: A resource book for teachers*. Essex, U.K.: Longman.
- Walker, C. (2009). The treatment of collocation by learners’ dictionaries, collocational dictionaries and dictionaries of business English. *International Journal of Lexicography*, 22, 281-299.
- Wang, J., Liang, S., & Ge, G. (2008). Establishment of a medical academic word list. *English for Specific Purposes*, 27, 442-458.
- Wray, A. (2000). Formulaic sequences in second language teaching: Principle and practice. *Applied Linguistics*, 21, 463-489.

CURRICULUM VITAE

Stephanie A. Balunda

EDUCATION

- M.A., English, Indiana University-Purdue University (IUPUI), Indianapolis
December 2009
- B.A., English & Comparative Literature, Indiana University, Bloomington
May 2007

CERTIFICATES

- Certificate in Teaching English as a Second Language
August 2009

PROFESSIONAL EXPERIENCE

09/2009 - 12/2009 **ESL Writing Instructor**

EAP/TOEFL Prep. for Malian Scholars, Indiana Center for Intercultural Communication, IUPUI

08/2009 **ESL Writing Instructor**

Tsuda English Program, Indiana Center for Intercultural Communication, IUPUI

06/2009 - present **Visiting Research Assistant**

Indiana Center for Intercultural Communication, IUPUI

08/2008 - 5/2009 **Associate Faculty**

English for Academic Purposes (EAP) program, IUPUI
Course taught: G011 ESL for Academic Purposes II

05/2008 - 05/2009 **Graduate Research Assistant**

Indiana Center for Intercultural Communication, IUPUI

12/2007 - 05/2009 **Substitute Teacher/ESL Tutor**

Atterbury Job Corps Adult Basic Education Center, Edinburgh, IN

02/2008 - 05/2008 **Practicum Student Teacher**

Afghan ESP and TOEFL Test Preparation Program, ICIC, IUPUI
Course taught: Academic/TOEFL speaking and listening

RELATED VOLUNTEER WORK

06/2009 - Present **ESL Tutor**
Atterbury Job Corps Adult Basic Education center, Edinburgh, IN.

06/2009 **Copy Editor**
Language Learning & Technology, University of Hawai'i at Manoa

02/2009 **Copy Editor**
Language Learning & Technology, University of Hawai'i at Manoa

CONFERENCES & PRESENTATIONS

42nd TESOL Annual Convention & Exhibit (New York, New York, April 2-5, 2008)
Attended only

4th Conference on Intercultural Rhetoric & Discourse (Indianapolis, Indiana, June 3-5, 2008)
Attended only

30th INTESOL Annual Conference (Carmel, Indiana, November 1, 2008)
Learning Academic Vocabulary: Strategies and Tools
Co-presenter: Suzan Stamper

30th ICAME Conference (Lancaster, UK, May 27-31, 2009)
Attended only

AWARDS

- Nominated "Favorite Professor" by student-athlete, IUPUI, 2009
- Nominated member of "21 Club", IUPUI, 2009

PROFESSIONAL ASSOCIATION MEMBERSHIPS

- Teachers of Speakers of Other Languages (TESOL), 2007-present
- Indiana Teachers of Speakers of Other Languages (INTESOL), 2009-present