

Winter 3-21-2014

# Opportunities for Incidental Acquisition of Academic Vocabulary from Teacher Speech in an English for Academic Purposes Classroom

Eric Dean Dodson  
*Portland State University*

Follow this and additional works at: [https://pdxscholar.library.pdx.edu/open\\_access\\_etds](https://pdxscholar.library.pdx.edu/open_access_etds)



Part of the [Applied Linguistics Commons](#), and the [First and Second Language Acquisition Commons](#)

**Let us know how access to this document benefits you.**

---

## Recommended Citation

Dodson, Eric Dean, "Opportunities for Incidental Acquisition of Academic Vocabulary from Teacher Speech in an English for Academic Purposes Classroom" (2014). *Dissertations and Theses*. Paper 1639. <https://doi.org/10.15760/etd.1638>

This Thesis is brought to you for free and open access. It has been accepted for inclusion in Dissertations and Theses by an authorized administrator of PDXScholar. Please contact us if we can make this document more accessible: [pdxscholar@pdx.edu](mailto:pdxscholar@pdx.edu).

Opportunities for Incidental Acquisition of Academic Vocabulary from Teacher  
Speech in an English for Academic Purposes Classroom

by

Eric Dean Dodson

A thesis submitted in partial fulfillment of the  
requirements for the degree of

Master of Arts  
in  
Teaching English to Speakers of Other Languages

Thesis Committee:  
Susan Conrad, Chair  
Kathryn Harris  
Darbra Smith

Portland State University  
2014

© 2014 Eric Dean Dodson



This work is licensed under the Creative Commons Attribution-NonCommercial 4.0 International License. To view a copy of this license, visit [http://creativecommons.org/licenses/by-nc/4.0/deed.en\\_US](http://creativecommons.org/licenses/by-nc/4.0/deed.en_US).

## **Abstract**

This study examines an English for Academic Purposes (EAP) teacher's speech throughout one curricular unit of an intermediate grammar and writing course in order to better understand which high-value vocabulary students might acquire through attending to the teacher and noticing words that are used.

Vocabulary acquisition is important for English for Academic Purposes students, given the vocabulary demands of academic language. The Academic Word List (Coxhead, 2000) has been shown to include important vocabulary in written academic texts, and has become a standard part of English for Academic Purposes curricula and pedagogical materials. Although explicit vocabulary instruction is important, research has shown that large amounts of vocabulary may be acquired incidentally by attending to meaning. Classroom instruction provides a great deal of input, and could potentially offer a chance for students to encounter and begin to learn academic vocabulary through incidental acquisition. However, existing research on incidental vocabulary acquisition in classrooms has focused on adult instruction and English as a Foreign Language settings, resulting in a lack of evidence about English for Academic Purposes classrooms.

To respond to these needs, this study analyzes the occurrence and repetition of Academic Word List items in the teacher's speech throughout two weeks of a course in an intensive academic English program in the United States. Two weeks of naturalistic class recordings from the Multimedia Adult Learner Corpus were transcribed and analyzed using the RANGE program to find the number of academic vocabulary types in the teacher's speech and how often they were repeated. Additionally, I derived categories of classroom topics and coded the transcribed speech in order to investigate the connection between topics and academic word use.

Academic Word List items are present in the teacher's speech, although they do not constitute a large proportion overall, only 2.8% of the running words. Most of the AWL types relate to specific classroom topics or routines. There are 13 AWL types repeated to a high degree, and 26 AWL types repeated to a moderate degree. These items are the most likely candidates for incidental vocabulary acquisition, though there is evidence from the videos that most of the students already understand their general meanings. It is unlikely that students could learn a great deal about AWL items that they were not already familiar with. However, it is possible that the teacher's speech provides incremental gains in AWL word knowledge.

These findings show that there may be a substantial number of AWL items that students learn about even before explicitly studying academic vocabulary. Teachers should try to draw out students' familiarity with these forms when explicitly teaching AWL vocabulary in order to connect familiar words with their academic meanings and uses.

## Acknowledgments

First, my thanks go to my committee for their precious time and feedback. I am also grateful to all those who worked on the Adult Lab School and Multimedia Adult English Learner Corpus, which made this research possible, and to John Hellermann, who first suggested that I use the set of IELP recordings therein.

Furthermore, this work would not have been possible without the support and forbearance of my family, friends, and L.B., who pushed me on countless times.

Finally, I would like to thank the teacher and students who agreed to have class recorded in full. Your decision to participate made this work possible.

Thank you.

## Table of Contents

Abstract.....	i
Acknowledgments.....	iv
List of Tables.....	vii
List of Figures.....	viii
Chapter One: Introduction.....	1
Glossary of terms.....	4
Chapter Two: Literature Review, Purpose, and Research Questions.....	6
2.1 The role of vocabulary in EAP settings.....	6
2.1.1 What do EAP students need to learn about vocabulary?.....	7
2.1.2 Importance of vocabulary in EAP.....	9
2.1.3 Importance of English academic vocabulary and the AWL.....	13
2.2 What is the role of incidental vocabulary acquisition in learning academic vocabulary?.....	20
2.3 Classroom discourse as an opportunity for vocabulary learning.....	26
2.4 Purpose.....	33
2.5 Research Questions.....	34
Chapter Three: Method.....	36
3.1 Overall Design.....	36
3.2 Context.....	38
3.3 Participants.....	40
3.4 Data Selection.....	42
3.5 Data Analysis.....	47
3.5.1 Transcription of Teacher’s Speech.....	47
3.5.2 Coding the teacher’s speech.....	52
3.5.3 RANGE Analysis.....	55



Chapter Four: Results and Discussion.....	59
4.1 Context: Class and category tokens.....	59
4.2 AWL Tokens, Types, and Families.....	64
4.3 Repetition of AWL items .....	71
4.4 AWL Occurrence by Classroom Topics.....	81
4.5 Further discussion: AWL use in context.....	87
4.6 Review of Findings .....	92
Chapter Five: Conclusion.....	95
5.1 Summary of Findings .....	95
5.2 Implications for EAP teaching .....	98
5.3 Relation to other research .....	102
5.4 Limitations .....	107
5.5 Future research.....	108
5.6 Conclusion .....	110
References .....	112
Appendix A: Institutional Review Board Approval Letter .....	118
Appendix B: AWL Type List from All Classes .....	119

## List of Tables

<b>Table 3.1:</b> Transcription comment abbreviations for activities that did not focus on the teacher’s speech .....	51
<b>Table 3.2:</b> Classroom Activity Categories .....	53
<b>Table 3.3:</b> Wordlists used by RANGE .....	56
<b>Table 4.1:</b> Number of tokens per word list per class .....	65
<b>Table 4.2:</b> Number of AWL Types and Families .....	66
<b>Table 4.3:</b> Repetitions of AWL types .....	72
<b>Table 4.4:</b> Normed frequency of the number of AWL types repeated six or more times / eleven or more times per 1,000 words .....	73
<b>Table 4.5:</b> AWL Types repeated 6+ times in a single class, ranked by total frequency .....	74
<b>Table 4.6:</b> Number of tokens per word list per classroom topic category .....	82
<b>Table 4.7:</b> Number of AWL types and families by classroom topic category .....	83
<b>Table 4.8:</b> Most frequent AWL Types by classroom topic category .....	84
<b>Table 5.1:</b> Comparison of AWL use in other classroom contexts .....	103

**List of Figures**

Figure 4.1: Tokens of teacher talk per class.....	61
Figure 4.2: Number of tokens per category per class.....	62

## Chapter One: Introduction

This thesis presents an exploratory analysis of the frequency of academic word list (AWL) items students were exposed to in the speech of their teacher in an English for Academic Purposes (EAP) grammar and writing classroom over one unit of the curriculum.

Gaining control over vocabulary in a new language can be a time-consuming, complex, and frustrating process for learners. Yet, as the prevalence of flashcards, vocabulary notebooks, software, and vocabulary quizzes suggests, it is an important part of studying English. Vocabulary learning is even more important for EAP students, as they need to learn and practice using a more specialized set of vocabulary used in academic settings. For these learners, vocabulary learning is crucial for their future academic success.

Coxhead's (2000) academic word list (AWL) presents a concrete set of word families that are frequently used in academic writing, which directly relates to the needs of EAP earners. As a result, the AWL has had a large impact on English instruction, and is widely used in EAP programs in classroom materials and curricular goals (for a review of the AWL's impact, see Coxhead, 2011). Students in many EAP programs study AWL items, either as part of their classes, or as self-study. However, the focused study of new words using

dictionaries, vocabulary notebooks, or other learning materials is just one part of vocabulary learning.

I have seen many students tackle the demands of academic vocabulary study, primarily through my experience tutoring students at Portland State University's Intensive English Language Program's Learning Center. Through tutoring, I have introduced many students to strategies and tools that fit their vocabulary learning needs and preferences. However, I also noticed that many students who began to study the AWL were already familiar with some of the words or word families in the list. Sometimes, students remembered seeing or hearing the words in situations that surprised me, such as television shows, host family interactions, or video games. Occasionally, it would take saying a word out loud to trigger a student's recognition. In some cases, an AWL word would seem familiar to students, even though they could not provide an accurate meaning. When this spark of recognition occurred several times, students seemed to grow more confident about learning AWL vocabulary, as if it were not far out of their reach after all.

My experience tutoring EAP students shows that explicit learning of vocabulary is not the only way that students are becoming more familiar with AWL items. Although learning about vocabulary in class and practicing items

with self-study methods is important for developing vocabulary knowledge, research has shown that vocabulary can be acquired incidentally in a variety of language modes when learners are focusing on meaning (e.g., Hulstijn 1992; Ellis, Tanaka, & Yamazaki 1994; File & Adams 2010). Some studies have investigated the opportunities for students to notice new words by attending to their teacher's speech, because attending to a language teacher's explanation of language structures or other content is itself a meaning-focused activity, and students spend a great deal of time in classrooms (Meara, Lightbown, & Halter, 1997; Lightbown, Meara, & Halter, 1998; Donzelli, 2007; Horst, 2009 and 2010). These studies have shown that ESL and EFL classrooms can provide some degree of repeated encounters with new vocabulary words, yet the evidence is mixed regarding the potential scale of this source of incidental vocabulary acquisition. Furthermore, there is a lack of research on this phenomenon in EAP classrooms, which are sites of learning for tens of thousands of students in preparatory programs around the world.

Thus, research is needed to address the exposure of EAP students to AWL vocabulary in the classroom. Also, this research should include an investigation of how likely it is that AWL items are noticed because of repetition, and what different types of classroom topics have to offer in terms of incidental

acquisition. I present an investigation of the opportunities for EAP students to notice AWL items in their teacher's speech throughout one unit of an intermediate-level grammar and writing course.

Chapter 2 includes a review of the literature on the importance of vocabulary and the AWL, incidental vocabulary acquisition, and classroom exposure to important vocabulary items. It also includes a section on the purpose of this thesis and the research questions. Chapter 3 presents the methodology I employed, including information regarding the context of the study, selection of the data, data transcription, coding, and analysis. Chapter 4 contains the results and discussion. Chapter 5 includes a summary of the findings, implications for EAP teaching and learning, connections to other research in the field, limitations, and suggestions for future research. This is followed by the list of references and appendices.

### **Glossary of terms**

The topic of vocabulary acquisition is associated with its own set of technical terms. Unfortunately, in writing about this topic, many authors do not define their terms, or even worse, use the term *word* in ambiguous ways. In order to make my own terminology clear, I provide the following glossary of terms in this thesis.

- incidental vocabulary acquisition: A gain in word knowledge that occurs while the language user is engaged in a task that is not directly related to learning vocabulary.
- lemma: a word and its inflections, limited to that word's grammatical category. For nouns, this includes the single and plural forms; for verbs, -ed and -ing participles; for adjectives, comparatives and superlatives.
- token: one instance of a word in a text.
- type: a word, limited to that specific sequence of letters, with no inflections. The narrowest definition of a word.
- running words: the total count of tokens in a text.
- word family: all the possible permutations of a base word (i.e., *headword*) combined with derivational affixes (e.g., un-, re-, over-, anti-, -tion, -ical, -ity, etc.). Typically contains multiple grammatical categories of words.

In addition to these more technical vocabulary, I use the following words in a specific manner:

- class: a single class period. For example, "In this class, the teacher gave students feedback on their first draft."
- course: the entire ten weeks of classes. For example, "In this course, the teacher gave students feedback on their drafts several times."



## **Chapter Two: Literature Review, Purpose, and Research Questions**

In this chapter, I review research that serves as important background for this thesis, and explain the purpose and research questions. The literature review contains sections on academic vocabulary in EAP programs, incidental vocabulary acquisition, and vocabulary exposure in ESL classroom discourse. This chapter concludes with the purpose and research questions of this thesis.

### **Literature Review**

Chapter one of this thesis introduced the theme of understanding the role of vocabulary in English for Academic Purposes (EAP) classrooms and programs. Accordingly, the literature review starts with a review of publications and research that help establish the importance of vocabulary for EAP students. In the second section of the literature review, I focus in on incidental vocabulary acquisition, a crucial yet under-studied part of vocabulary learning. Finally, I build upon these sections to make the argument that incidental vocabulary learning in EAP classrooms may play a vital role in students' vocabulary learning.

#### **2.1 The role of vocabulary in EAP settings**

Three topics are important for understanding the role of vocabulary in EAP pedagogy: what vocabulary knowledge EAP students may need, the

importance of that vocabulary for students in an EAP program, and the use of the Academic Word List as a target for English learners.

### **2.1.1 What do EAP students need to learn about vocabulary?**

Students who want to prepare for using English in academic environments must learn a great deal about both general high-frequency English vocabulary and more specialized academic words and meanings. Before gauging which words are crucial and how many of them there are, I review works from second language acquisition regarding the construct of word knowledge.

First, it is important to establish what it means to learn or know a word. In the past, research on vocabulary in the fields of SLA and education has focused on acquiring form-meaning connections (especially written word form to either an L1 translation or a definition in the L2) due to historical traditions of testing vocabulary learning through multiple choice or definition-based tests (see Nation 2001 for overview). However, modern research into L2 vocabulary acquisition suggests that this testing metaphor vastly over-simplifies the incremental nature of word knowledge learning (Schmitt, 2010). Pedagogical perspectives that reduce vocabulary acquisition to only learning word definitions overlook other types of word knowledge. For example, a student might be able to match a word with its definition in a written test and at the same time, unable

to use that word in a productive sentence, identify its part of speech, or identify its common collocates. Such a student could not be said to have learned that word well, or at least not to its full extent.

Instead of a binary distinction where a word is either *known* or *not known*, many experts in second language vocabulary acquisition favor the construct of word knowledge, which has many components that may be learned in an incremental fashion. Aspects of word knowledge include a word's phonological form, spelling, connections to concepts, synonymy/antonymy, affixes, grammatical features, and sociolinguistic use. Nation (2001) explains that these aspects of word knowledge can be grouped into three categories: lexical form (spelling, pronunciation, recognition of a form, word parts), meaning (translation to L1, connection to other L2 words, related words), and use (part of speech, collocational patterns, associations with genre or formality). Each of these categories includes procedural knowledge that aids both the understanding and production of language, known as *receptive* and *productive* knowledge. Finally, word knowledge incorporates the mode of language (i.e., spoken or written). These different categories of word knowledge do not exist separately, but unite in the learners' knowledge of a given word in terms of what control they have over the word.

This word knowledge perspective, as supported by Nation and others (e.g., Nagy, Herman, & Anderson, 1985; Schmitt & Meara, 1997; Webb 2007), portrays the complexity of learning vocabulary, and the many pieces of information that are required to understand and use a word accurately and fluently. Furthermore, each type of word knowledge exists on a gradient of no or little control, to a precise, accurate level of control. It is unlikely that learners gain all aspects of word knowledge at once while simultaneously going from no knowledge to perfect control over a word (see Schmitt, 2010 p. 20-21 for overview). Learners improve their vocabulary through incremental gains in word knowledge (and the types of word knowledge they possess), as shown in a number of studies designed to measure and analyze these incremental gains (e.g., Nagy, Herman, & Anderson, 1985; Webb & Kagimoto, 2009; Milton, 2009). In addition studies suggest that learners may acquire vocabulary knowledge in small amounts over time, instead of all at once as a result of focused study. I return to the topic of vocabulary acquisition in section 2.2, on the role of incidental vocabulary acquisition.

### **2.1.2 Importance of vocabulary in EAP**

Learning vocabulary in English is a large and complex task, even considering non-specialized or everyday language. Nation (2001) estimates that

adult native English speakers have a vocabulary of about 20,000 word families, although English learners may not necessarily need to learn that amount, depending on their goals. Instead, if the task of vocabulary learning is to understand most written and spoken language (i.e., the goal for many language learners), a more modest estimate is 8,000-9,000 word families for written language, and 6,000-7,000 word families for spoken language (Nation, 2006). It should be noted that both of these estimates are for receptive vocabulary only (i.e., reading or listening). However, even this estimate of receptive vocabulary represents a formidable challenge for language learners, as discussed further in section 2.2.

Although learning vocabulary may seem like an imposing task, it is nonetheless correlated with overall language skill (Staehr, 2008; Nation, 2006). Reading is a fundamental language skill for students who will study in English-language universities, and vocabulary has been shown to be a vital part of reading comprehension in studies such as Laufer (1989), which provided evidence that in order for a reader to comprehend a text, at least 95% of the total words must be known. In the years after Laufer's research (and perhaps due to its influence), many other researchers in the field of reading, vocabulary, and second language learning have conducted similar studies (e.g., Laufer &

Ravenhorst-Kalovski, 2010; Hsueh-Chao & Nation, 2000). These studies have replicated or come close to Laufer's initial figure. As such, it seems that the consensus in second language research is that readers must know anywhere from 95-98% of the running words of a written text in order to understand it. Schmitt, Jiang, and Grabe (2011) provide strong evidence for these figures, though they add that there is no specific threshold that makes comprehension possible; still, they argue that readers of academic texts should aim for understanding 98% of the running words. Altogether, these studies contradict the notion that English learners can overcome gaps in their vocabulary knowledge by guessing meanings from context or other reading strategies.

Although most research has focused on the role of vocabulary in reading comprehension, it is clear from research and practice that it plays an important role in listening comprehension as well (Goh 1999; Bonk, 2000; Nation, 2006; Lynch, 2011). Kelly (1991) provides evidence of "lexical ignorance" as substantially detrimental to listening comprehension among advanced L2 users, and even claims that it is "the main obstacle to listening comprehension with advanced foreign language learners" (p. 135). Rost (1994) shows that non-native speakers' unfamiliarity with key terms in an academic lecture was associated with difficulty interpreting the lecturer's definitions, examples, and the

difference between the two. In discussing the implications for pedagogy, Rost recommends “building a lexical base for lecture instruction” and “generating lists of key terms for students to prepare prior to a lecture” in order to improve listening comprehension (p. 113).

Vocabulary is also crucial for performance on high-stakes English language tests, an important step in applying for university study. Milton (2009) found that vocabulary proficiency tests correlated significantly with students’ IELTS scores across all skill sections, accounting for as much as 60% of the variation in writing scores, and 50% of the variation in reading (p. 179). The importance of vocabulary for standard tests was also shown by Staehr (2008), who studied the correlation of scores between language skill sections on a national EFL test; vocabulary test scores were significantly correlated with writing and reading scores, and, to a lesser extent, with listening scores. This shows that learners’ general level of vocabulary knowledge is correlated with success in meeting institutional admission requirements necessary for academic study.

The importance of vocabulary is perhaps the greatest for non-native English users enrolled in English language universities. The role of vocabulary in reading and listening becomes even more important once students are attending

English language lectures, reading large amounts of course materials, and other activities that require mastery of a wide range of vocabulary. There is also the issue of what professors and content instructors expect from students. In a survey of 234 lecturers in tertiary institutions, Ferris and Tagg (1996) found that respondents across all disciplines were concerned about L2 English speaker students' reading and familiarity with technical vocabulary. This suggests that content lecturers are aware of their students' vocabulary use, and expect EAP programs to prepare students for instruction in English.

Vocabulary learning is also part of EAP and English for Specific Purposes (ESP) students. Elisha-Primo et al (2010) showed that graduate students in an Israeli EAP program ranked the importance of vocabulary significantly higher than grammar, reading, or writing. Similarly, in a needs analysis of students in a Polish English institute, Mazejko (2011) found 70% of the students indicated that they wanted to spend more time on vocabulary development, and that vocabulary had the highest mean rank of importance.

### **2.1.3 Importance of English academic vocabulary and the Academic Word**

#### **List**

Topics in this section include: what vocabulary items are the most helpful for preparing learners for English academic environments, the development of



the Academic Word List (Coxhead, 2000), and its subsequent use in EAP programs.

In general, a small number of words (types or families) will account for a large percent of the total words found in a given sample of English, while a large number of words are used very infrequently (Nation, 2001; Nation, 2006; Schmitt, 2010). For English teachers and learners, the relative frequency of a word provides one principle for deciding whether it is worthwhile to spend time and resources on learning a given word. There have been a number of attempts to construct word lists based on word frequency in a large number of texts in order to provide specific word-learning goals for material design and language pedagogy. For example, the General Service List (GSL) (West, 1953) aimed to include the 2,000 most frequent word families in English, and has been used in countless texts and courses that have aimed to teach the most common words in English. Nation (2001) reports that the GSL covered 90% of the words in a corpus of conversations, 80% of a newspaper corpus, and 78% of an academic text corpus (p.17). This shows that the GSL is still relevant in studying language, though there have been many attempts to update or improve upon it, including a recent re-formulation conducted by Brezina and Gablasova (2013).

As the GSL has served as a resource for students, teachers, and researchers interested in general English vocabulary, there has also been a need for word lists that focus on general academic vocabulary. For some time, the most popular academic word list in the realm of applied linguistics was the University Word List (UWL), a compilation of several lists of academic vocabulary drawn together by Xue and Nation (1984). However, this list was surpassed in terms of quality and popularity by the Academic Word List (AWL), compiled by Coxhead (2000).

The AWL has had a tremendous impact on applied linguistics and EAP pedagogy, in large part due to its design and the underlying principles of the corpus used to create it. Coxhead assembled a corpus of 3.5 million running words from a range of representative types of written texts (including textbooks, journal articles, and laboratory manuals) from multiple fields. Moreover, Coxhead included four main branches of disciplines (arts, commerce, law, science), each with sub-genres (for example, science included texts from biology, chemistry, mathematics, and physics). These principles provided a general representation of academic language that would not be influenced too heavily by any one discipline, ensuring that the list would be valuable for EAP students interested in a wide range of subjects.

In order to compile the actual AWL from this corpus, Coxhead extracted a word list from the academic corpus and then applied a set of criteria to make sure that the AWL had words that were generally useful for academic contexts. She excluded items that occurred on West's GSL in order to eliminate general (not academic in meaning or use) high-frequency words from the list. The final AWL included word families that occurred 100 times or more, and only included items that met certain range criteria. Specifically, each word family had to occur at least 10 times in each discipline, and at least once across 15 or more of the subject areas (p. 221). This means that the AWL contains useful word families for students in a variety of fields, while limiting the number of subject-specific items. The result of applying the criteria above was the AWL, with 570 word families comprising 3,110 types. Example AWL types include *analysis*, *concept*, and *data*.

Coxhead's analysis of AWL occurrence in two corpora of academic texts showed that the items were indeed used more frequently in academic writing. The AWL accounted for 10% of the running words in Coxhead's academic corpus, and 8.5% of another corpus of academic texts used to test its robustness. For both corpora, the GSL covered a majority of the running words (academic corpus: 71%, second corpus: 66%). Finally, in order to show that the AWL truly represented academic genres, and not overall high-frequency words, it was

compared with a corpus of fiction texts. In this analysis, the AWL only covered 1.4% of the tokens, which showed that it was relevant for academic language in particular.

In the years after its publication, many studies have duplicated Coxhead's results and provided further evidence of the AWL's usefulness in capturing important academic language (see Coxhead 2011 for an overview). A Google Scholar entry for the original 2000 study and publication shows that it has been cited in over 1,100 articles and books. Furthermore, many English language material developers have embraced the AWL, and even refer to it in EAP textbook titles, such as *Focus on Vocabulary: Mastering the Academic Word List* by Schmitt and Schmitt (2005). In addition, the AWL has become a part of EAP program curricula, such as the AWL-driven vocabulary assessment goals in the Portland State University Intensive English Language Program (IELP) high-intermediate and advanced levels of study (2012). Some simple internet searches reveal that the AWL is integrated into many other intensive English programs' curricula, such as the University of Pennsylvania (2013), Texas A&M University (2007), and University of Lincoln, U.K. (2013). This means that students in EAP programs like the PSU IELP have to learn AWL items not only for their future

success, but also in order to succeed in their EAP studies (e.g., if they want to understand their textbooks or do well on classroom assessments).

There has also been criticism of the AWL and attempts to update or replace it, which provides more evidence that the AWL continues to play an important role in the field. Hyland & Tse (2007) argue that it over-simplifies the complexity of academic sub-fields, and that EAP instruction should include more exposure to subject or genre-specific vocabulary. Others have argued that EAP programs need to go beyond the AWL because it does not include enough items to truly prepare L2 speakers for academic contexts (e.g., Schmitt & Schmitt, 2012).

A new academic vocabulary list has recently appeared. Gardner and Davies (2013) compiled an Academic Vocabulary List (AVL), which they argue improves upon the AWL. The AVL departs from the AWL in several ways: it is a list of lemmas (not families), and it does not exclude GSL items. This is crucial, the authors argue, because many high-frequency words from general use also have alternate meanings and usage in academic context. Although Gardner and Davies employed several design principles to limit this over-inclusion, it could be the case that this decision over-compensates and includes items that should be considered high-frequency words. Setting aside the design decisions of the lists,

Gardner and Davies show that, compared to the AWL, the AVL covers a greater portion of the academic sub-section of the Corpus of Contemporary American English (COCA) (Davies, 2008-).

Despite these criticisms and the development of new word lists, the AWL's prominence in the field of EAP pedagogy and applied linguistics remains unmatched at this time. It seems likely that the AWL will continue to be relevant for EAP pedagogy for the near future because of its prevalence in curricula and pedagogical materials. For example, the EAP program studied in this thesis has no plans at this time to revise their curriculum to change from the AWL to the AVL. Regardless of future developments in EAP and applied linguistics, it is clear that the AWL is a useful, empirically-constructed tool in understanding academic language.

To sum up, there have been many word lists designed to focus on academic vocabulary, including the UWL, AWL, and AVL. In particular, the AWL has been widely used by researchers, textbook creators, and EAP programs alike, due to the principles of its design. Also, further research using the AWL (as in Coxhead, 2011) has provided additional evidence that it contains important vocabulary for written academic genres. In closing, Coxhead (2000) points to the need for students to encounter AWL words in many ways:

Direct teaching through vocabulary exercises, teacher explanation, and awareness raising, and deliberate learning using word cards need to be balanced with opportunities to meet the vocabulary in message-focused reading and listening and to use the vocabulary in speaking and writing. (p. 228)

## **2.2 What is the role of incidental vocabulary acquisition in learning academic vocabulary?**

This section of the literature review presents research on the conditions necessary for learners to acquire word knowledge, with a particular focus on incidental acquisition. Ultimately, the goal of this section is to present the background necessary to understand how EAP students may (or may not) learn academic vocabulary, and what role incidental acquisition may play. However, as I show below, there is a lack of concrete evidence regarding the scope of incidental vocabulary acquisition in classrooms, which comprise a large amount of linguistic input.

Past developments in second language acquisition help us understand what conditions EAP programs need to bring about in order for students to learn vocabulary. As teaching methods and research in second language acquisition have shifted from the audio-lingual method to include communicative

competence and comprehension, more varied perspectives and approaches have been used to understand vocabulary learning. In describing the lexical component of his “i+1” theory, Krashen (1982) wrote that “It may be the case that if we supply enough comprehensible input, vocabulary acquisition will take care of itself” (p.81). One central assumption of this theory was that by attending to meaning, language learners are able to acquire vocabulary, even if they are not consciously attempting to.

As part of an increased interest in comprehensible input, second language acquisition researchers began to use the construct of incidental learning, defined as the acquisition of lexis when the direction or goal of a task is not to learn words, but to understand a passage or attend to another language feature. Hulstijn (2007), in a review of the concept, writes, “telling or not telling students that they will be tested afterwards on their knowledge is the critical operational feature distinguishing incidental from intentional learning” (p. 267-268). Instead of a sole focus on repetition as a driving factor in vocabulary acquisition, researchers in the field began to look at the role of negotiation and learner output typified by the interactional hypothesis espoused by Long (1981).

Incidental vocabulary acquisition is consistent with Krashen’s theories, and it completes part of the puzzle of how vocabulary is acquired. Nation (2001)



argues that with the considerable amount of vocabulary and types of word knowledge that language learners need to gain, explicit instruction and definition of vocabulary items are inefficient and unlikely ways for learners to make large gains. Instead, he argues, incidental acquisition is responsible for a large part of vocabulary growth, advancing at a relatively slow, steady pace.

On the topic of attention and vocabulary learning, Ellis (1994) writes that attention alone may be sufficient for acquiring new forms, but attention and a greater deal of cognitive processing may be needed in conjunction to acquire a form-meaning connection. Laufer and Hulstijn (2001) review past studies and theorize that depth of processing (i.e., the amount of mental effort that learners have to exert to complete a given task) leads to superior gains in incidental vocabulary acquisition. Their model of task involvement includes a need (there is an obvious gap in form/meaning that needs to be filled), search (the action of trying to locate or remember a lexical item) and evaluation (assessing whether the item is appropriate given the meaning, form, or context of the language structure). Keating (2008) provides further confirmation for this task involvement model.

Studies of incidental acquisition have shown that repetition, comprehension, and noticing lead to gains in word knowledge (Hulstijn, 1992;

Zahar, Spada, and Cobb 2001; Webb, 2007). In their review of second language vocabulary acquisition studies, Zahar, Spada, and Cobb (2001) write that previous studies have found as few as 6 and as many as 16 repetitions in a short period of time are necessary for some degree of word learning to take place among a majority of learners. However, as Webb (2007) points out, there have yet to be enough studies (or any one study big enough) to create a convincing rule of thumb in terms of the number of repetitions it takes to learn a word. Furthermore, given the role of context and attention, there is likely no one number of repetitions that will lead to acquisition among all learners. Lastly, this research has almost entirely focused on encounters with vocabulary items in written language, ignoring opportunities in spoken language.

Listening also presents an avenue for incidental acquisition, yet there is a lack of research on this topic. The situation is summed up by Schmitt (2010):

There has been much less research on spoken discourse, again probably because it is harder to research. Therefore, there is a big gap in the field's understanding of how spoken discourse and vocabulary. We have little idea of how vocabulary is learned from listening, how many repetitions it requires, or what makes a word salient for learning in spoken discourse. (p. 38)

Although there is not a large amount of research on the topic of incidental vocabulary acquisition from spoken language, Vidal (2003) provides strong evidence that spoken language can lead to gains in vocabulary through incidental acquisition. Ellis (1994) postulates that “It is possible that a single exposure is sufficient for some learning to take place... Learners, for example, may be able to recognize it as a “real” word where previously they could not” (p.11).

Joe (2010) shows how spoken and written language combine in an EAP setting and lead to incidental vocabulary acquisition . The author meticulously recorded one learner’s vocabulary encounters in class and in assigned texts throughout one term, and presents evidence that the number of repetitions of key vocabulary items was the key to the student recalling words during interviews or testing. These repetitions of vocabulary included instances where the teacher was not drawing the student’s attention to it. Regarding the quantity of encounters with a word, this study provided evidence that 4 encounters with a vocabulary item was not enough to add to the learner’s level of word knowledge. A familiarity with the word form only was associated with a mean of 7.5 repetitions, and a concrete gain in a previously unknown word’s productive use was associated an average of 18 repetitions. These results show the strong

correlation between repetition and incidental vocabulary acquisition, this time in an EAP setting.

Apart from this case study, it is unclear how much repetition occurs in the linguistic input of EAP students. There is scant evidence regarding how often EAP students see or hear important AWL words, understand their meanings, or notice the way they are used. Studies such as Miller (2011) have investigated opportunities to encounter AWL items in written form, but there has been a lack of studies on opportunities to hear them in classroom settings. This is important because of the substantial number of AWL items there are to learn and the learning burden associated with them. If there is a relatively small amount of exposure to AWL vocabulary in classroom settings, then explicit teaching and learning of word associations may indeed be necessary. On the other hand, if students are given many opportunities to see and hear collocations in meaningful contexts, they may acquire those associations incidentally, and this may help students when they read academic texts or are tested on AWL items later in their EAP studies. This could potentially be an important part of the way that EAP students learn word knowledge, as students in intensive English programs spend a great deal of time in classes attending to the speech of their

teacher and classmates. In order to understand this issue, it is necessary to review research conducted in English learning settings.

### **2.3 Classroom discourse as an opportunity for vocabulary learning**

Only a few studies have investigated the opportunities for vocabulary acquisition in classroom discourse, with a focus on the lexical profile of teacher talk. More often, research has focused on explicit vocabulary instruction, as in Folse (2010), or on the opportunities for incidental acquisition in other types of listening (e.g., Webb & Rodgers, 2009). Less research has focused on opportunities for incidental vocabulary acquisition in classroom discourse.

Meara, Lightbown and Halter (1997) applied tools that had been used to analyze vocabulary in texts to a small corpus of classroom transcripts in order to investigate incidental vocabulary learning opportunities in English classrooms. They note that the teacher's speech presents a large amount of language input, which means that students may have opportunities to attend to meaning, notice new words, recall them, and gather contextual, collocational information about their use. This is exactly the kind of attention discussed by Laufer and Hulstijn (2001) in their task-based involvement model of vocabulary acquisition. Meara, Lightbown and Halter studied ten 30-minute samples of class instruction from a primary school ESL program for French speakers in Quebec that employed a

communicative approach. They found that the patterns of word frequency and total counts were remarkably uniform across teachers. Furthermore, the researchers focused on “unusual words” (words that did not occur in the most common 2,500 words of Nation’s word lists), and found that there were an average 2.75 “unusual words” per 500 words. Though this first seemed to be a low amount, the authors noted that the students’ daily exposure could be as high as 50 unusual words every day, extrapolating from the patterns in the data. They judged this amount, equivalent to 250 words per week, to be a fairly rich exposure to words that students (likely) don’t already know.

The authors of the above study continued their line of inquiry with a study that included the same setting, but expanded the data set to include elementary classrooms in Quebec taught with audio-lingual methods, and adult English courses taught with communicative methods (Lightbown, Meara, & Halter, 1998). In this study, the authors compared the number and distribution of lemmas between 19 different teachers in different settings. Although this study predated the creation of the AWL, it makes use of Nation’s UWL list, which offers a less precise measure of academic vocabulary. The mean number of UWL lemmas in 1500-word samples of the transcripts was fairly low. For example, in the adult English courses, the experienced teachers ranged from an average of 4.0

to 7.5 UWL items per 1500 words, which shows that variation in lexical characteristics of teacher talk between speakers falls within a relatively limited range. In comparing the lexical characteristics across classroom settings, the authors found that the intensive communicative elementary setting teachers used significantly more low-frequency lemmas than the other settings. Finally, the authors turned towards an analysis of word repetition in the data, and found that teachers in all contexts were remarkably similar in the number of words they repeated, though this included words that learners likely knew already.

In revisiting the above studies, Horst (2009) continued to expand the methodology of analyzing classroom vocabulary to include a deeper investigation into repeated items. The data from this study came from the same elementary EFL setting as the two studies by Meara, Lightbown, and Halter, and adapted their methodology. Horst provided evidence that as many as 20 word unknown word families a day might be made available for uptake through teacher talk. Whereas previous studies had difficulty operationalizing words that were likely unknown to learners, Horst made use of a corpus of student writing to identify a set of words that learners likely had control over, and counted opportunities to learn words in the teachers' speech that did not occur frequently in the writing corpus.

Perhaps because previous studies were limited to small samples of classroom teaching from a number of teachers, Horst (2010) studied one classroom in depth through one full term. In this study, Horst measured the frequency of repetition of unusual words in an adult ESL classroom in Quebec over all 36 hours of the teacher's speech. In addition to investigating the occurrence of unusual words, Horst considered six repetitions of a given word family in more than one transcript as enough for students to retain form-meaning knowledge of that item, using the minimum number suggested in the review of lexical acquisition conducted by Zahar, Cobb and Spada (2001). Of the 949 word families above the 2,000 word list level throughout the entire class, only 245 of them were repeated more than six times, which equated to an average 14 word families per class. Horst writes that even though it is likely that students could learn these 14 word families per class, this was a meager amount of vocabulary learning considering the students' needs.

In both studies, Horst (2009, 2010) found that the exposure to unusual words (i.e., not likely known by students, operationalized by off-list words) in teacher talk occurred in different amounts according to the categorization of the teacher's speech by topic and function. This methodological innovation showed that teacher talk was not homogenous, but varied according to classroom topics.



In Horst (2009), the transcripts of teacher talk had been coded using a modified COLT (Communicative Orientation of Language Teaching) observation scheme (Spada, N., & Fröhlich, M., 1995). The data were categorized as “classroom procedures, language-related episodes..., written texts presented aloud, text-related discussion and meaningful interaction,” (Horst 2009, p.56) although only classroom management, reading texts aloud, and language-related episodes were frequent enough to compare with each other. Horst hypothesized that pre-written texts would have the highest percentage of off-list families. However, Horst found that language-focused talk had the highest rate of off-list families (135 families in a 3000 word sample). Pre-written texts read aloud and classroom management were lower, and nearly the same (87 and 84 families in a 3000 word sample).

Horst (2010) follows the same methodology, with similar results that suggest language-focused talk featured the most exposure to unknown words. However, since this research, there has not been further investigation in the variation of vocabulary exposure in classrooms as it relates to classroom topics and routines. As a result, it is not clear whether Horst’s findings are similar to other classroom settings, or if the COLT categories capture variation in classroom topic well.

Other studies have also measured overall exposure to important vocabulary in classroom settings, but with slightly different methods or research questions. Donzelli (2007) analyzed the teacher's oral utterances over an entire school year (55 hours of instruction) in a primary school EFL class. The author found that the teacher's utterances featured an average 24 words above the 2,500 most common word level (again, using Nation's 1986 word lists). She claims that this qualifies as a satisfactory amount of vocabulary exposure, especially as it exceeds the amount of exposure found in class texts. Donzelli also tested the retention of vocabulary items, and found that the mean test scores on words repeated 10 times or more were significantly higher than words repeated less often (p.120). This supports the claim that classroom discourse may provide support by not only creating opportunities for noticing, but by aiding retention.

In contrast, a more pessimistic outlook on the potential for teacher talk to provide vocabulary learning support was given by Tang (2011), who used yet another measure of vocabulary richness. Tang calculated vocabulary range and type-token ratios of six EFL teachers in one week of instruction, and compared it to the vocabulary-learning goals in that institution. Tang concluded that none of the courses provided rich vocabulary instruction and exposure to low-frequency words because only 22% of the word types used were on the UWL or off-list. In

addition, many of those off-list types were proper names. Although it isn't clear what Tang would have considered to be "adequate," she concludes that "the classroom did not seem to provide a lot of opportunities for an 'i+1' condition" (p. 48).

To sum up studies on classroom vocabulary exposure, the findings are inconsistent, and leave many questions unanswered regarding the lexical characteristics of teacher talk. On one hand, studies of some elementary EFL classrooms that employ communicative language teaching methods suggest that teacher talk could provide enough rich input in order for the incidental acquisition of a large amount of vocabulary over time (Meara, Lightbown, and Halter 1997; Lightbown, Meara, and Halter, 1998; Horst, 2009; Donzelli, 2007). In contrast, Horst's follow-up study (2010) showed that few words were being repeated enough to lead to meaningful gains in vocabulary. Furthermore, Tang (2011) provided more evidence that teacher talk does not provide substantial vocabulary support in adult ESL and university EFL classrooms. At this time, there has been no attempt to replicate an analysis of spoken vocabulary use in an intensive EAP course.

## 2.4 Purpose

This literature review has made a case for the importance of vocabulary learning in English for Academic Purposes (EAP) programs, including the special role of the Academic Word List (AWL) in preparing learners for university settings. Although explicit instruction is important for developing vocabulary knowledge, research has shown that vocabulary may be acquired incidentally when learners are engaged in various tasks. Incidental acquisition of vocabulary also depends on noticing new information, which means that prominence through repetition in a short period of time or other means can increase the likelihood that something about a given word is learned, even if it is a low-level awareness of the existence of that word. This kind of incidental learning is an important channel for easing the burden of learning AWL items, as intensive EAP students spend a considerable amount of time in classrooms attending to the teacher's language.

There is some evidence that there may be consistent exposure to new vocabulary through teacher talk in some contexts, and that different classroom topics and activities present different amounts of exposure (Horst, 2009; Horst, 2010). However, previous research is inconclusive about how much vocabulary is available for students to attend to in classroom settings. Furthermore, existing

research focuses on adult instruction and EFL settings, resulting in a dearth of evidence about EAP classrooms. Research is needed to address the opportunities for EAP students to learn AWL vocabulary in the classroom through listening to their teacher's speech.

The purpose of this thesis is to understand English learners' exposure to AWL words by listening to their teacher's speech in an EAP classroom throughout one curricular unit of the course. The findings will help identify what role (if any) incidental acquisition might play in EAP students learning AWL items, even at intermediate skill levels. It could be the case that students who take high-intermediate or advanced-level courses that require them to learn AWL items come into those courses with substantial amounts of AWL knowledge already. Understanding this potential source of knowledge better could lead to enhancements in EAP curricula, improvements to the way that AWL items are taught, and a better understanding of how EAP students rise to the challenge of learning academic vocabulary.

## **2.5 Research Questions**

In order to understand what AWL items EAP students might learn from their teacher's speech, this study will address the following research questions:

1. How many AWL types and families are students exposed to in the teacher's speech throughout one curricular unit in an EAP classroom?
2. How frequently are AWL types repeated in the teacher's speech in each class?
3. What classroom topics in the teacher's speech feature the highest frequency of AWL tokens and types?

The answers to these questions will be applied in a discussion of the likelihood of incidental acquisition of academic vocabulary from an EAP teacher's speech.

## **Chapter Three: Method**

The sections in this chapter elaborate on the research design I employed in order to answer my research questions. Here, I describe: the overall design of this thesis, the context of the study, participants, data collection, and data analysis. I conclude the methods chapter with a discussion of the reliability of this thesis.

### **3.1 Overall Design**

This study is an exploratory analysis of the exposure to AWL items that all the students in an EAP classroom received from the teacher's spoken language throughout one curricular unit. Specifically, I transcribed the teacher's speech during teacher-focused portions of class and quantified the frequency of AWL types, the number of AWL tokens, and repetition of AWL types in each class. I selected eight contiguous classes that constituted a curricular unit in order to describe the variation in total amounts of teacher talk, AWL use, and classroom topics across multiple classes, as the students made progress through the curriculum. In addition to the transcription and quantification of AWL use, I coded the teacher's speech according to categories of classroom topics in order to account for variation of AWL use across classroom activities and topics.

This study uses naturalistic data from an existing ESL class, which means that it offers a detailed look at a real classroom, and not an experimental instructional setting. Furthermore, it means that the teacher's language behavior was not influenced by the design of this research. Though the teacher and students were aware of the recording equipment, the course followed the same curriculum and procedures as other courses in the program. Moreover, there was no way that the teacher would have guessed that her spontaneous speech would be the focus of this study, as these recordings were made before my research process had begun. Because this data reflects typical class procedures for this program, the results may be of value for educators in other EAP programs who see similarities between their classrooms and this one.

The data for this study came from one teacher's speech, which allowed me to understand details of one course as it progressed over time. Though the teacher's individual speech characteristics may have influenced certain lexical features of her speech, past research on the lexical features of English language teacher speech has shown that teachers in similar teaching settings are remarkably similar in terms of their speech's lexical frequency profiles (in particular, see Horst 2009, p. 57 for a comparison of word family use among



teachers, and Lightbown, Meara, & Halter 1998, p. 329 for a comparison of lemma use among teachers).

In order to address the research question regarding classroom topics, I coded the teacher's speech according to topical categories similar to those described by Horst (2010), which were adapted from the COLT instrument (Spada & Fröhlich, 1995). Horst's research on an adult English language course in Montreal showed that language-focused talk had a higher frequency of likely unknown lexis than classroom management talk, which was relatively simple. Because of these findings, I decided that types of teacher talk and classroom language were potential moderating variables on the frequency of AWL use in teacher talk. However, I found that the COLT instrument was not well-suited for describing this classroom and teacher. Instead of forcing the data to fit into the COLT scheme, I noted common classroom routines and topics, and created an ad hoc coding scheme based on emergent categories from the data. Finally, I coded all of the transcripts according to these category criteria, which are described in the section regarding Data Analysis.

### **3.2 Context**

The context of this study is the intensive English language program (IELP) at Portland State University, located in Portland, Oregon. The IELP has six levels

of instruction designed to prepare students for academic study at the undergraduate or graduate level. The course levels range from entry-level low-beginner (“Level E”) to advanced (“Level 5”). At each level, courses are organized by language skill (i.e., Speaking/Listening, Reading, Grammar/Writing, and an elective course). Full-time students have 18 hours of class every week. Though each course is focused on certain language skills, a wide variety of teaching methods are used. Class size ranges from 12-18 students. Teachers in this program have a graduate degree in TESOL or a related field such as education or communication and often have many years of experience teaching ESOL. The classroom context itself, including the reasons for its selection, are described below in the section on Data Collection.

The data for this study was recorded in 2008 as part of the Adult Lab School project at Portland State University, in a classroom equipped with six video cameras and four sets of microphones in order to record naturalistic English language instruction. The Adult Lab School project was supported in part by grant R309B6002 from the Institute for Education Science, U.S. Department of Education, to the National Center for the Study of Adult Learning and Literacy (NCSALL). The Adult Lab School project data was accessible as part of the Multimedia Adult English Learner Corpus (MAELC) (see Reder, Harris,

and Setzler, 2003 for an overview of the MAELC). Although most of the courses recorded in the MAELC “Lab School” classrooms were taught as part of an adult ESL program, several IELP courses were also recorded there, including this Grammar/Writing course.

### **3.3 Participants**

The data for this study was a pre-recorded IELP course from the Multimedia Adult English Learner Corpus (MAELC). All the teachers and students in the classrooms gave their informed consent to be recorded for the purpose of research in applied linguistics. This informed consent was obtained from all participants before the recordings were added to the MAELC. In compliance with the policies of Portland State University, I obtained approval from the Portland State University Human Subjects Research Review Committee to use data from the MAELC for this study (see Appendix A for approval letter from PSU HSRRC).

In watching the videos and observing student and teacher behavior, I treated them with the respect they should be granted, in accordance with the MAELC Confidentiality Agreement and User License. I refrained from making critical or evaluative statements about their performance or appearance in the class. Finally, I watched and transcribed classroom recordings in settings where

other people could not see the video player nor hear what was said by the teacher and students in this class.

The participant of this study was the instructor of an IELP intermediate-level combined grammar and writing course. Like all teachers in the IELP, this teacher had a graduate degree in TESOL or a related subject and substantial teaching experience. The teacher in this study was thus generally representative of current teaching standards in this and similar EAP programs. Furthermore, the teacher had taught the same subject and level at least once before, which reduced potential variation due to inexperience with the curricular goals. In this study, I specifically avoided critiquing the teacher's pedagogical approach or any aspect of her teaching behavior. On the contrary, I considered these recordings as a record of an exemplary EAP teacher who held to the high professional and academic standards of the IELP and the larger community of ESOL education. Moreover, I considered the frequency of high-value vocabulary in classroom speech to be language behavior co-constructed by the teacher, students, classroom context, course materials, and curricular goals.

The students were an important part of classroom discourse, but I did not include their speech in transcriptions. Even though I did not transcribe the

students' speech, I listened to their questions and comments in order to understand the context of the teacher's speech.

### **3.4 Data Selection**

I used the Spring 2008 intermediate level combined Grammar/Writing course from the MAELC as the main source of data for this study. The recording procedures employed meant that the classroom video and audio quality was quite good, which was immensely helpful for the observation and transcription process described in further detail below.

I selected this course in particular because of its level and topic. An intermediate level course was appropriate for this study because this level typically features fairly complex language, yet does not have an explicit focus on academic vocabulary, which could obscure the role of incidental acquisition.

One of the reasons for this thesis is to investigate the exposure that students have to AWL items before they begin to focus exclusively on academic language or directly study AWL items. Beginning-level courses usually feature simplified or otherwise limited language that does not likely include many instances of AWL items, which makes them inappropriate for this kind of exploratory study. At the other end of the spectrum, advanced-level courses often incorporate an explicit focus on studying AWL items, which would make it

difficult to categorize items in the teacher's speech as opportunities for incidental acquisition.

In addition, grammar and writing courses in the IELP are not characterized as focusing on vocabulary development, which means that the course topic was also ideal for investigating opportunities for incidental (rather than explicit) vocabulary learning. Because students in the course attended to the teacher for meaning, and only very rarely in order to learn vocabulary, the course fit the conditions for incidental vocabulary learning. In sum, the level, topic, and setting of this high-intermediate Grammar/Writing course meant that it was an appropriate choice for the research questions at hand.

In order to become more familiar with the course and to choose a curricular unit for transcription, I watched the entire term's worth of recordings of this classroom. In this survey of classes, I took notes regarding the class topics, activities, assignments, common routines, and other notable features in the class. For some classes, I watched the entire two hours; for others, especially after I grew attuned to the general rhythm of the teacher and students, I watched 20-30 minutes of the class. In those cases, I skipped some portions or fast-forwarded through activities, and focused on smaller portions of time throughout the class in order to understand the lesson. In many cases, I re-watched classes in order to

add to or refine my initial observations. These notes helped me grasp the underlying context of the class routines and relationships as they developed throughout the term, and they proved to be useful in coding the categories of teacher speech later on in this study.

In my initial observations of this course, I noticed that class time was generally split between teacher-led activities or instruction and group or individual tasks. The teacher-led instruction included both grammar and writing topics, though more time in-class was spent on writing, because students were expected to spend time outside of class working on a self-guided grammar portfolio project. Generally, the teacher's speech consisted of giving instructions for activities or homework, explaining language or pertinent concepts, eliciting answers from the class, or giving tests. Almost every class session featured relatively long periods (20-30 minutes at a stretch) of individual or group work time, where students worked on activities such as writing tasks, grammar discovery exercises, journal writing, or group quizzes. This division between teacher-focused class time and individual or group work was clearly important in terms of my research questions, because only the former presented an opportunity for all students to attend to the teacher's speech. I discuss how I addressed this division further below, in the sub-section on transcription.

The course met on Mondays, Wednesdays, and Fridays for two hours and ten minutes per class. Generally, the class proceeded without a break during the first three weeks, though in the fourth week, the teacher began to give students a ten minute break in the middle of the class, after noticing their low energy after the one hour mark. The first week of the course was not included in the multimedia corpus. Attendance records were not provided along with the video, but from my observations, there were occasional absences and only a few instances of students joining the class late. Attendance was never below 11 students, and never exceeded 16.

After watching the entirety of this course, I selected eight classes to transcribe. In selecting these classes, my criteria were that they should feature a clear progression of the curriculum regarding specific skills or tasks, cover approximately two weeks, and not include cancelled classes, holidays, or instances of substitute teaching. By progression of the curriculum, I mean a series of lessons that begin with the introduction and explanation of a topic; controlled and guided practice with the given skill, task, or concept; and various activities or other supporting lessons that lead to extension activities related to the skill or task. It was important to me to select contiguous classes in order to see how the teacher's language varied from class to class as topics were introduced and



developed. I applied these criteria by reviewing my observation notes and re-watching certain classes to confirm my choice.

The eight classes that I chose met the criteria above because they had a clear coherence as a curricular unit and did not include any cancellations, holidays, or substitute teachers. The selection of classes began with the Friday of the second week of the term, and ended on the Monday of the fourth week. These classes clearly fit the description of a curricular unit given above. Specifically, they form a unit on writing a cause and effect essay. The teacher took a process writing approach, and led students throughout this unit on a sequence of explanations and exercises that helped them build up to writing outlines, first drafts, written peer reviews, second drafts, and ultimately, their final draft. This shows the progression from teacher-led instruction to increasing levels of student control, culminating in individual writing. The selection of classes includes all the lessons related to cause and effect essay writing, from the introduction of concepts and form, to the teacher's last in-class feedback and guidance on their final drafts. These details show that these eight classes fit the criteria above and were appropriate for selection and further study.

I chose these classes because of the clear unit on writing cause and effect essays, but this selection included a progression of grammar instruction and

activities as well. The sequence of grammar topics did not occur in parallel with the writing lessons, but because of the lesser role of grammar during class time, this fact did not affect my decision to select classes that formed a writing unit. This selection included several lessons regarding verb tense formation, and, in the eighth class, a lesson introducing passive formation. These grammar lessons, however, were less common and received less in-class time. In fact, the most important grammar-related portion of this course was an independent grammar portfolio that students worked on outside of class. Accordingly, the teacher wove grammar lessons throughout the course, with some topics spanning several weeks. As a result, despite the fact that I chose these classes based on the focus on essay writing, grammar was an auxiliary topic during class time.

### **3.5 Data Analysis**

The procedures I followed included transcribing the teacher's speech, coding the transcripts for activity types, preparing text files for analysis, and using software to calculate the number and frequency of AWL types and families. Each of these steps is described below.

#### **3.5.1 Transcription of Teacher's Speech**

I transcribed the teacher's speech by typing in a plain-text document while simultaneously watching and listening to a class using the MAELC ToolBox

software program. I included a short header in each text file for recording the class date and time. During transcription, I generally watched the video at half speed and typed in the document simultaneously. I used a broad orthographic transcription that included long pauses (measured in seconds) and interruptions, but did not incorporate other fine-grained aspects of speech. I did not mark lengthened syllables, micropauses, or other elements common to narrow transcriptions. After every 10 to 15 minutes of transcription, I re-read the transcript to ensure its fidelity and to correct any typing errors. Once I completed a transcript for a class session, I reviewed the document again, this time while playing the video recording.

All headers, comments, and codes were added between angle brackets (i.e., < and >), which were ignored by the software used to analyze the lexical features of the text. I marked pauses or hesitations that were less than one second with individual hyphens. For longer pauses, I wrote the length in seconds followed by the letter *s*, all in angle brackets (e.g., <2s>). In cases when the pause in speech was due to the teacher writing on the erasable board or projector transparency, I typed <writes>. Although I did not transcribe students' speech, I noted when students spoke by using the code <ss>. I did not transcribe proper

names of the participants, using <sname> for students' names, and <tname> for the teacher's.

In the transcripts, I included limited comments to note the video time, and physical context of the classroom when they were helpful for understanding the transcription. In this excerpt, for example, I noted that the teacher handed out papers with prompts in order to make it clear what the teacher is referring to:

so you can choose any method either clustering listing or outlining whatever works for you <4s>  
<begins to hand out prompts> ok you guys are going to have these so as soon as you finish taking notes separate all of those so that you can see all the prompts

One instructional feature that I noted in comments was explicit vocabulary teaching. I used the code <EV> every time the teacher provided the meaning or other aspect of word knowledge regarding a specific word. I noted these moments because they provide another way that the teacher raised students' awareness of vocabulary items. In this excerpt, for example, the teacher is reviewing terms in a class reading, and elicits the meaning of *logger*:

what's a <EV> logger <ss> yeah a person who cuts down trees

In order to address the question of what vocabulary students in this course were exposed to, I limited my transcription of teacher speech to portions of the class that were teacher-centered or where the teacher was clearly

addressing the entire class. I did not include speech that was recorded during group or individual activities, even when the teacher was giving feedback or answering questions with individual students during those times. I did not include this speech because at those times, the other students were not attending to the teacher's speech. This means that only one student was actually being exposed to that language, while the others were communicating with each other or engaged in other tasks and language inputs. Including this language would misrepresent what all the students heard in the class. However, I included any speech where the teacher gave follow-up instructions to groups during group activity when she gave similar instructions one by one to groups (i.e., when all groups heard nearly the same instructions). In these cases, I transcribed one instance of the explanations/instructions.

Although I did not transcribe group or individual activities, I recorded the amount of time spent on these activities. In many of the classes, the teacher-focused portions made up a majority of the class time. Although group and individual activities accounted for a majority of the entire class time over the eight classes, the teacher-centered portions were vital to the nature of the course, as this was the time that the teacher set the agenda, assigned homework, instructed students on grammar and writing, reviewed student answers from

activities, and gave instructions for group or individual tasks. The different types of task were very clear, based on the teacher's instructions and my observations of the entire course.

Table 3.1 shows the transcription comment abbreviations that I used to note different types of classroom activities which did not involve attending to the teacher's speech. When the focus in a class shifted away from the teacher for these types of activities or for a class break, I typed the corresponding abbreviation and noted the time that the activity started and ended. I did this while doing my first transcription of a class. Afterwards, I recorded the time spent on each non-teacher-focused activity. This allowed me to tally the amount of time spent attending to the teacher and her speech compared to other tasks. These results are reported in section 4.1 of Chapter Four.

Table 3.1: Transcription comment abbreviations for activities that did not focus on the teacher's speech	
Abbreviation	Meaning
<gg>	Group grammar tasks
<pg>	Pair-based grammar tasks
<ig>	Individual grammar tasks
<gw>	Group writing tasks
<pw>	Pair-based writing tasks
<ip>	Individual writing tasks
<go>	Other group tasks
<break>	Class break time

### 3.5.2 Coding the teacher's speech

Once the transcripts were compiled, I developed categories of classroom topics based on re-occurring classroom activities and types of course content. The model for this procedure can be found in work by Horst (2010), who adapted the categories of classroom management, language-focused talk, personal anecdotes, and pre-recorded or pre-written speech from the COLT instrument (Spada & Fröhlich, 1995). However, unlike Horst, I defined categories based on the specific context of this class, and not based on an external model. The advantage, though, was that the categories used in this study are more meaningful and relevant for the class being described. In addition, the final categories used in this study could be considered more detailed than the COLT codes. However, the ad hoc categories carry the disadvantage of not being directly comparable to other studies, including Horst's, that use the COLT. Table 3.2 displays the category names, abbreviations, definitions, and examples.

Pre-written text, though it did not align to a specific topic, constituted a category because it was not true spoken language, but written language read out loud, which distinguished it from all the other categories. This category, however, still represented part of the oral input that students attended to in the class. In order to accurately assign this code, I made use of the classroom video to

Table 3.2: Classroom Activity Categories		
Category	Code	Definition
<b>Activity Instructions</b>	IN	Procedure related to an in-class language-learning task or guidance for the task mid-way through you can do this on your own <hands out papers to class> fill in each um biography with the correct form of the verb
<b>Assignments</b>	HW	Assignment of homework, guidance for what to study out of class, description of grammar portfolio. ok so I've given you one more sheet of paper about Jack Kerouac and Gary Snyder and it's right here <2s> this one is homework
<b>Other Class Management</b>	CM	Agenda setting; comments related to course objectives; distributing papers; requests related to discipline; talk about writing conferences. if I can I'd like to take the brainstorming notes that you did in your group <1s> please write the names of your group members on those papers
<b>Language focus: Grammar</b>	LG	Language-focused talk related to grammar, explicit instruction of grammar, review of grammar points. there's a difference in English between things that we think of as points in time <writes> and things that we think of as taking place <writes> over a period of time
<b>Language focus: Writing</b>	LW	Language-focused talk related to writing, explicit instruction, discussion or review of writing in homework or in-class tasks. the structure of the cause and effect essay is just like what you did in the classification essay and just like what is shown on the paper about the basic structure of the five paragraph essay
<b>Anecdotes</b>	AN	Personal anecdotes or comments that do not relate to classroom topics there's a joke about uh what is it this is a novel that I read a while ago and they were making fun of um teachers and policemen...
<b>Quizzes</b>	TE	Teacher giving past verb tests (i.e., reading lists of verbs for students to inflect) buy verb number three choose <1s> I choose to learn French choose <5s> verb number four come <1s> will you come to my birthday party
<b>Pre-written text</b>	PW	Instances where the teacher was reading out loud from texts that were written in advance. Amy has been learning English since she was in middle school ... the students had been waiting for over an hour before tname arrived



make sure that this was only used for times when the teacher was reading directly from a written text. I did not include repetitions or elaborations of the text when the teacher was clearly not reading from a text. After defining the categories of classroom topics, I coded the entirety of the transcripts based on the content of the language, the physical context, and my understanding of this course's classroom routines. The goal was not to assign each clause constituent to a category, but to capture shifts from one topic or classroom routine to another. For that reason, the minimum length of a coded portion of text was two clauses. This ensured that instead of coding individual words or statements, I categorized shifts in classroom discourse (i.e., language beyond the clausal level).

Several factors aided the coding decisions. First, the teacher was well-organized and often made explicit transitions to new topics or activities. Once the teacher started on a given topic or activity, she usually stayed on that topic until its conclusion, without switching to other topics abruptly. Also, my knowledge of this course's routines and curriculum helped me understand how the teacher's speech fit into the categories of classroom topics. For example, my familiarity with the curriculum and grammar-related outcomes of the course helped distinguish grammar-focused language from writing-focused language. By the time I coded the data, I had already watched these classes at least two

times, and as many as five times. As a result, I was very familiar with the teacher's routines in general and the sequence of topics in each class in particular. All of these elements helped provide empirical reasons for coding decisions. Generally, I found the categories and coding decisions unambiguous. In a few cases where coding decisions were more ambiguous, an additional reading of the transcript and surrounding context was enough to clear up any initial ambiguity. Because of the procedures and rationale above, I did not perform a check of inter-rater reliability.

I coded the transcripts using Microsoft Word's style tools. For each category, I assigned a style and corresponding background color. This allowed me to easily select all of each category throughout the transcript, and save it in its own text file. After all the above, I prepared the transcripts for lexical frequency analysis by reviewing each text file using Microsoft Word's spell-check features, to ensure that I had not made any typos or included inconsistent spellings.

### **3.5.3 RANGE Analysis**

I used the RANGE program (Heatley, Nation, & Coxhead, 2002) to analyze the text files. RANGE divides the words in a written text into those that belong to West's General Service List (GSL) of the first 1000 most frequent words in English, the GSL list of the second thousand most frequent words, the AWL,

and “off-list,” the words that do not belong to any of the above word lists. It offers detailed information about AWL items, reporting the number of families, types, tokens, and type-token ratio. It also provides the number of tokens that each word list covers and the total number of tokens in the text. The RANGE wordlists are summarized in Table 3.3.

<b>RANGE Wordlist Name</b>	<b>Description</b>
<b>1k</b>	The first 1,000 most common word families in English (as compiled by Nation et. al)
<b>2k</b>	The second 1,000 most common word families in English
<b>AWL</b>	The Academic Word List (Coxhead, 2000)
<b>Off-list</b>	Types that do not occur on any of the above lists

I analyzed the eight class transcripts in order to address research questions two and three, regarding the use and repetition of AWL tokens, types, and families. In order to address research question three, regarding the use of AWL items in each category of classroom topics, I ran RANGE on the aggregate text files for each category. For each processed transcript I analyzed with RANGE, I recorded the following information in a spreadsheet:

- Total word count
- # AWL families
- #AWL types

- #AWL tokens
- #AWL tokens as % of running words
- AWL types and frequency count
- # of repetitions of AWL types and families
- % of off-list running words

The final spreadsheet showed all the above information for each class, which allowed me to see classes side-by-side. In addition to the quantitative results, I also saved the RANGE list of AWL types used in the transcripts in spreadsheet form.

This final spreadsheet contained the quantitative results necessary to answer research question one. In addition, I spent some time looking through the list of AWL types in order to understand any patterns in the items that were spoken by the teacher.

Regarding research question two, I calculated the greatest number of repetitions of each AWL type in a single class based on its raw count. I examined the distribution of the data and categorized them into four different levels of repetition: single use of a type, low (2-5 repetitions), moderate (6-10 repetitions), and high (11 or more repetitions). These categories are adaptations from Horst (2010): I modified the low and moderate cutoff points based on the distribution of the data. Also, I counted the number of AWL types in each repetition category in each class. I then normed this figure (per 1,000 words) in order to compare the

frequency of repetition across classes that featured varying amounts of teacher talk. After calculating the quantitative figures, I re-examined the AWL types that occurred in each level of repetition.

In order to address research question three, regarding the features of AWL use in each category of classroom topics, I looked for similarities and differences in the number of AWL types and families in each category, specifically focusing on which categories featured the most AWL types. I also sorted the list of AWL types in order to see what types were the most common in each category, and what similarities and differences those categories had.

Finally, I used concordancing software to better understand patterns and contexts of AWL use in the data, and to provide illustrative examples for discussion. I did this with only a few of the most frequent items. In addition to searching for these examples in the data, I ran concordance searches for them in the Contemporary Corpus of American English (Davies, 2008-), in order to generally compare the use and contexts of AWL items in these data with academic writing.

## **Chapter Four: Results and Discussion**

In this chapter, I explain the results of the analyses and discuss them point by point. This chapter includes sections on the context of the results; AWL tokens, types, and families; repetition; and categories of classroom topics. Each section contains results and discussion. A final section discusses the contexts of the AWL items in this classroom compared to their use in academic writing.

### **4.1 Context: Class and category tokens**

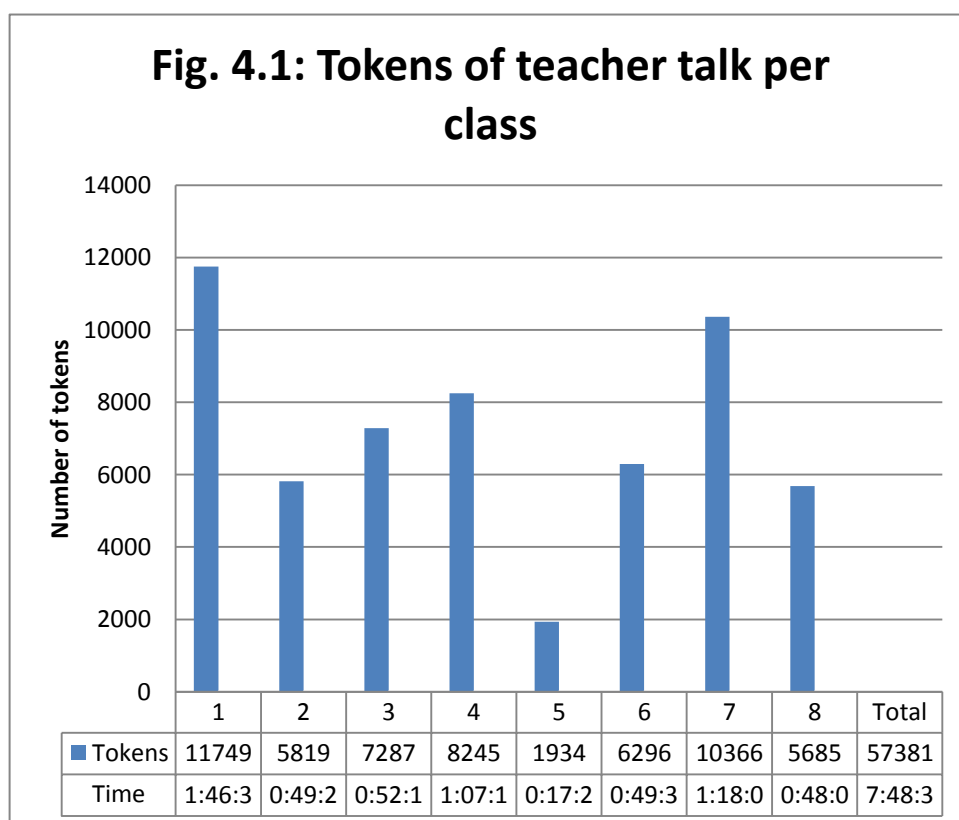
In this section, I first describe the data in terms of the total tokens of teacher speech in each class session and the breakdown of classroom topic categories, which provides background and context for the results. This information, although not the focus of this thesis, is important in establishing a context for the results that follow. This context aids the interpretation of results in terms of classroom routines, activities, and topics, and not simply AWL use on a class-by-class basis. Indeed, the variation in activity time, amount of teacher speech, and classroom topics shows that each class is a unique mix of these factors, and not an undifferentiated phenomenon.

First, it should be noted the number of tokens per class varies across the eight classes. Figure 4.1 below shows the number of tokens of teacher speech

during teacher-focused activities in each class. The total number of tokens in the data is 57,381. However, the amount of teacher speech is not evenly distributed. Figure 4.1 shows the fluctuation in the amount of teacher talk, from a minimum of 1,934 tokens in class five to a maximum of 11,749 tokens in class one. This means that any quantitative results regarding the average number of AWL items per class must be understood in terms of the varying amount of teacher talk. Results from a class with a large amount of teacher speech, such as class one or seven, should not be directly compared with classes with less teacher speech, such as class five or eight. Instead, it will be important to report and compare the normalized rate of AWL use and repetition.

The variation in the number of tokens per class can be better understood by considering the amount of teacher-focused time, which is also listed in the data table in figure 4.1. The total amount of teacher-focused time is 7 hours, 48 minutes, 31 seconds, which comprises 43.9% of the total time over eight classes. Teacher-focused time per class varies from as little as 17 minutes, 22 seconds (or less than 15% of the total class period) in class 5, to as much as 1 hour, 46 minutes, 25 seconds (or nearly 80% of the total class period) in class one. This helps to explain the variation in the amount of teacher talk. The class with the lowest number of tokens, class five, was nearly completely devoted to group

projects and individual writing as a “work day.” This variation shows that opportunities for students to learn vocabulary from the teacher’s speech were distributed unevenly, and that class time included a balance of teacher-centered instruction and group or individual activities.

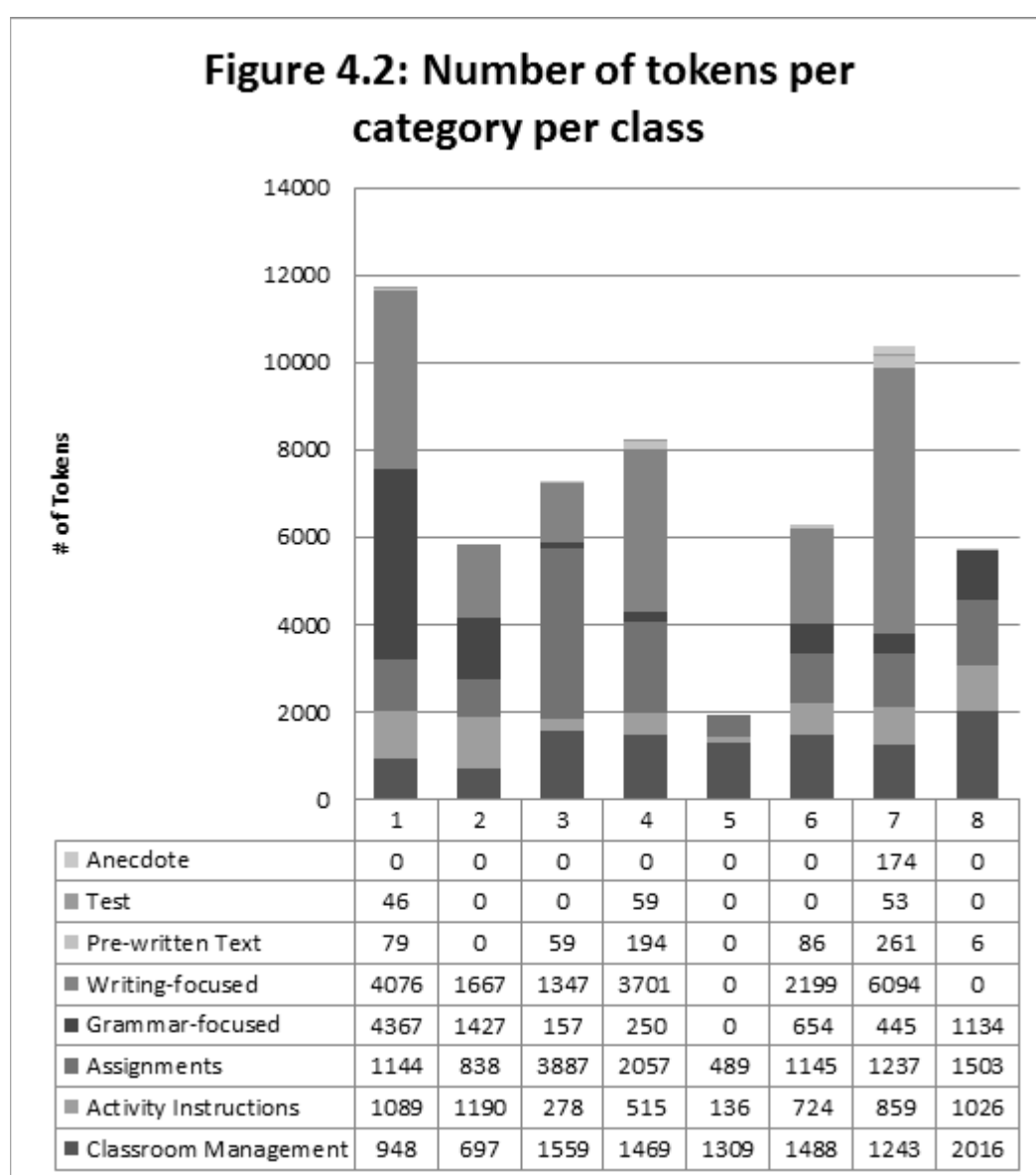


Finally, the variation between classes shows the importance of the decision to focus on a contiguous set of classes instead of sampling single classes from throughout the term. A sampling paradigm that did not include contiguous classes could obscure the kind of variation that is on display in the data here.

In addition to the variation in tokens and teacher-focused time, the



distribution of classroom topic categories varied across classes as well. Figure 4.2 displays the number of tokens in each category for each class. This provides a bird's-eye view of what the teacher talked about, which provides useful context for understanding common topics. Each category shows considerable variation across classes, with the exception of anecdotes, pre-written texts, and tests, which



remain very low or zero throughout. For example, Writing-focused language has 6,094 tokens in class seven, but disappears completely in class five and eight.

Interestingly, three categories occur in all classes: classroom management, activity Instruction, and assignments.

Finally, this context is vital for understanding sections the analysis of the occurrence of AWL items in categories of classroom topics (section 4.4). Similar to the variation across classes, it is not feasible to compare the categories directly because of their size differences. Writing-focused language comprises the largest amount of the total data, with 19,084 tokens (33%), followed by assignments, with 12,300 tokens (22%), classroom management, with 10,729 tokens (19%), grammar-focused language, with 8,434 tokens (15%), and activity instructions, with 5,817 tokens (10%). The categories of pre-written text, tests, and anecdotes make up less than 1% of the total tokens each. As with the class totals of teacher talk, it was necessary to normalize the rate of AWL use before comparing these categories.

## 4.2 AWL Tokens, Types, and Families

### *Results*

In this section, I report the number of AWL tokens per class, the number of AWL tokens as a percentage of the total tokens per class, the number of AWL types per class, and the totals of these measures for the entire data. Along with the AWL, I report the number of tokens the other word lists provided by RANGE (i.e., 1,000 highest-frequency word families, second thousand highest-frequency word families, and the remaining tokens that were not on any of the previous lists), because these figures give context for the AWL tokens, types, and families in each class.

Table 4.1 shows the number of tokens in each of the RANGE program's wordlists. The row with AWL results is in bold. The table shows that the AWL covers from 1.8% to 3.6% of the tokens in each class, and an overall figure of 2.8% of the entirety of the data, or 1,634 instances of AWL families out of 57,381 tokens of speech. This makes it the least frequently used of the four word lists overall, and on a class-by-class basis, with the exception of classes four, six, and seven, where it has a higher frequency than the off-list category. Unsurprisingly, the 1,000 highest-frequency word family list is by far the largest category overall, followed by the 2k word list and the off-list.

In addition to the number of AWL family tokens that occur in the data, RANGE provided the number of AWL types and families in each class and across all classes. These results are summarized in Table 4.2. Note that because many of the types and families are repeated across classes, the total

Class	1	2	3	4	5	6	7	8	Com- bined
<b>AWL Tokens (%)</b>	<b>212 (1.8%)</b>	<b>150 (2.6%)</b>	<b>234 (3.2%)</b>	<b>250 (3%)</b>	<b>54 (2.8%)</b>	<b>228 (3.6%)</b>	<b>346 (3.3%)</b>	<b>160 (2.8%)</b>	<b>1634 (2.8%)</b>
<b>1k List Tokens (%)</b>	10591 (90.1%)	5276 (90.7%)	6516 (89.4%)	7399 (89.7%)	1753 (90.6%)	5650 (89.7%)	9368 (90.4%)	5166 (90.9%)	51719 (90.1%)
<b>2k List Tokens (%)</b>	443 (3.8%)	221 (3.8%)	291 (4%)	394 (4.8%)	61 (3.2%)	261 (4.1%)	353 (3.4%)	184 (3.2%)	2208 (3.8%)
<b>Off- List Tokens (%)</b>	503 (4.3%)	172 (3%)	246 (3.4%)	202 (2.4%)	66 (3.4%)	157 (2.5%)	299 (2.9%)	175 (3.1%)	1820 (3.2%)
<b>Total Tokens</b>	11749	5819	7287	8245	1934	6296	10366	5685	57381

column is not equivalent to the numerical sum of the eight classes (i.e., it is a count of the total number of types and families used throughout the entire data). The number of AWL types in each class varies from 31 to 132 types, with a total of 334 types used throughout the entire data. The number of AWL families varies from 24 to 101, with a total of 213 families used throughout the entire data.

Table 4.2: Number of AWL Types and Families									
Class	1	2	3	4	5	6	7	8	Combined
AWL Types	110	70	88	100	31	81	132	59	334
AWL Families	84	54	70	76	24	66	101	49	213

The entire list of AWL types, ranked by their frequency in the data, can be found in Appendix B.

### *Discussion*

In this section I describe some of the trends that appear in the way that students were exposed to AWL tokens, types, and families. I provide examples of AWL use in context to show that generally, the occurrences of AWL items is not necessarily conspicuous, but fits the needs of classroom instruction and the content of the course. However, there is mixed evidence whether the AWL items in this course are semantically or grammatically similar to their use in academic content.

By looking at the AWL types, tokens, and families that students were exposed to, we can identify some items that appear to be closely linked to classroom topics, and discover that some items are likely already known (or assumed to be known by the teacher). A general trend is that many of the AWL types are related to the context of a grammar and writing class, and are used in a

way that seems natural given the functions and topics involved. Some key examples of high-frequency AWL types that fit this description are *thesis*, *conclusion*, and *individual* (in the sense of individual work compared with group work). Indeed, it would be hard to imagine that a teacher would be able to avoid using these words in conducting a grammar and writing course. These points are supported by looking at the use of these examples of AWL types in context.

For example, AWL types have been underlined in the excerpt below, from class seven, during the teacher's instruction on the use of repetition and variety in essay writing:

in your conclusion you need to have your - you need to have your main idea and you need to have your supporting ideas which basically means you need your thesis statement and so some students every term some student goes to their thesis statement says aha I have that right here <typing noise> copy paste ah finished no <laughs> not finished because yes you do need to repeat yourself but you've got to use different words to do it

This excerpt contains two instances of *thesis*, both in the context of a thesis statement, and one instance of *conclusion*, which are used in many other classes (*thesis* occurs in seven of the eight classes and *conclusion* occurs in three). In the above excerpt, the teacher is establishing or reaffirming her expectation that student writing should contain a conclusion that features a paraphrase of their thesis statement. Because of the strong link between class content and *thesis* and

*conclusion*, these AWL types are among the most frequently used in the data. This is a clear example of how some AWL items play a central role in the teacher's classroom language due to their connection to the course content.

Many of the AWL types, including the examples above, are terms that are used to refer to features of a text. This accounts for both their occurrence in a writing course and in the AWL, which was compiled from written sources. There are many such textual references used by the teacher in this course, including *paragraph*, *topic*, *edit*, *section*, *draft*, and *author*.

The use of *individual*, unlike the examples given above, may not be as closely related to the content of the course, but is clearly related to classroom procedures. The following excerpt comes from the second class; the teacher is explaining how students should make use of their group collaboration for an outline assignment:

ok this is a nonflexible requirement of this assignment so make sure that you do this with the group <writes> as an individual you'll do the rest of your outline so you'll find <5s> <grabs paper from front desk> the outline looks like this ok <1s> this part here your thesis statement your topic sentences - um all of those are the same everything else on the outline should be slightly different for each person

In this excerpt, the teacher uses *individual* to differentiate from what students should do as a group. Here, the use of *individual* stems not from the classroom

content of grammar or writing, but of the typical procedure of giving instructions and explaining what should be done as groups and as individuals.

An interesting feature of this excerpt is that the teacher elaborates on these instructions and the concept of *individual*, saying "... the outline should be slightly different for each person." At least in this instance, it seems that students who didn't comprehend the word *individual* (or any other part of the explanation) were given additional support through this elaboration. Though it is possible that this also provided students with a synonym for the term *individual*, there are four clauses that occur between the AWL term and the elaboration, containing over thirty words of speech and a long pause. Given the processing demands of aural cognition, it seems unlikely that students attending to meaning would make the connection between *individual* and its elaboration, though it is not completely impossible.

The examples given thus far show another pattern: the prevalence of nouns, especially in the most frequently-used types. This seems to correlate with the content-specific AWL types, such as *thesis*, *paragraph*, *topic*, *tense*, and *transition*, as well as the types related to assigning homework or tasks, such as *draft*, *journal*, *team*, *conference*, and *project*. Many of these nouns are unavoidable considering the topic of the course.



Regarding what students might be able to learn from exposure, there are many AWL types that occur in the data that students likely understand already because they are commonly used in many English language instruction contexts, even in beginning levels, or because the forms themselves occur with relatively high frequency in spoken English. In this particular EAP program, it is difficult to imagine a teacher conducting class without using terms such as *assignment*, even in beginning-level courses. Similarly, it would be expected that students who had taken a beginning EAP writing course would recognize terms like *conclusion*, *topic*, and *draft*. It does not seem likely that the students in this course had never noticed or used these words before. Although it is possible that students still gain incremental word knowledge and control over these types because they notice them in new collocational or grammatical contexts, my intuition as a teacher is that many beginning-level students would have at least some receptive knowledge of many of the words in the Appendix.

Finally, the AWL types in the teacher's speech are almost completely unrelated to the few instances of explicit vocabulary instruction in these classes. Explicit instruction here means interactions where the teacher called specific attention to the word and either provided the meaning, elicited it from students, or provided lists of important words to know. For example, the excerpt below is

from the teacher reading off the overhead projector during part of a lesson on using transition words:

after all - as an illustration - for example -  
 for instance specifically - to take a case in  
 point xx there are slight variations in what kind  
 of information you might put after which of these  
 words but all of them would be used to introduce  
 to your reader hey reader here comes an example

These interactions were not frequent, nor were they part of the classroom routines. Only six AWL types received explicit attention regarding their meaning: *function, illustration, instance, relevance, specifically, and whereas*.

Considering that the entirety of the classes featured 334 AWL types, it is clear that explicit vocabulary instruction is linked only to a small number of AWL types that students were exposed to. However, I return to the topic of explicit vocabulary focus in the repetition discussion (section 4.3) to discuss interesting connections between explicit vocabulary focus and repetition.

### **4.3 Repetition of AWL items**

#### *Results*

The next set of results addresses the repetition of AWL types in each class. Table 4.3 shows the level of repetitions of AWL types in each class and across all classes in four categories: single occurrences (exactly one instance of a type with

no repetition), low repetition (two to five instances), moderate repetition (six to ten instances), and high repetition (eleven or more instances). The single-instance category has the highest frequency: 176 AWL types occur no more than one time in any of the classes. The next category, low repetition, includes 119 AWL types. Most importantly in terms of existing research on the importance of repetition, the moderate and high repetition categories include 26 and 13 AWL types, respectively.

In addition to the total count of repetitions, it is important to find the relative frequency of high AWL type repetition due to the variation in amount of teacher talk between classes. As section 4.1 showed, the number of tokens varied from class to class. It follows that classes with more tokens might have more

	1	2	3	4	5	6	7	8	Combined
<b>Single</b>	63	46	47	54	21	42	71	40	176
<b>Low repetition (2-5 Reps)</b>	42	19	31	38	9	29	50	11	119
<b>Moderate repetition (6-10 Reps)</b>	5	3	6	4	1	8	6	6	26
<b>High repetition (11+ Reps)</b>	0	2	4	4	0	2	5	2	13

repetition in part because of the larger amount of teacher speech. In order to compare the classes without this distortion, I calculated the frequency of each

category as a normed rate. Table 4.4 below displays the normed frequency of selected categories of repetition per thousand words. This shows normalized figures for how many AWL types occur at a moderate or higher level of repetition, and how many occur at a high level of repetition. The normed frequency of moderate and high repetition AWL types ranges from a low of 0.43 types per thousand words in class one, to a high of 1.59 types per thousand words in class six. The figures for the normed frequency of only high repetition types are, naturally, lower. They range from zero for classes one and five, to a high of 0.55 AWL types with high repetition per thousand words in class three.

<b>Class session</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>Combined</b>
<b># AWL types: Moderate or High Repetition per 1000 tokens</b>	0.43	0.86	1.37	0.97	0.52	1.59	1.06	1.41	0.68
<b># AWL types: high repetition per 1000 tokens</b>	-	0.34	0.55	0.49	-	0.32	0.48	0.35	0.23

Table 4.5 provides the AWL types from the moderate and high categories (i.e., the combination of AWL Types repeated 6-10 times and 11+ times). This includes the number of classes they occurred in (their range), their total frequency in the data, and the count of their repetition from the class in which they were used the most times, ranked by total frequency. The types in bold are

in the same family with at least one other type on the list. For example, number four *topic* is in the same family as number 26 *topics*. There are fifteen such types on this list that are related to each other; they account for seven distinct families.

Table 4.5: AWL Types repeated 6+ times in a single class,  
ranked by total frequency (family members in bold typeface)

Type	Range	Total Freq	Highest # of rep.	Type	Range	Total Freq	Highest # of rep.
1. thesis	7	59	17	21. contrast	4	18	10
2. passive	6	53	32	22. <b>draft</b>	6	18	8
3. paragraph	6	52	18	23. process	6	17	6
4. <b>topic</b>	6	51	20	24. job	5	16	6
5. <b>assignment</b>	7	46	16	25. relevance	2	15	9
6. <b>editing</b>	7	36	8	26. <b>topics</b>	4	15	8
7. tense	6	35	13	27. team	4	14	6
8. transition	3	30	25	28. project	3	13	6
9. journal	6	28	11	29. <b>edit</b>	4	12	9
10. <b>conclusion</b>	3	27	23	30. parallel	2	12	8
11. <b>link</b>	4	25	14	31. <b>conclusions</b>	2	11	10
12. <b>assignments</b>	5	23	7	32. clause	3	11	8
13. <b>linking</b>	3	21	17	33. <b>conference</b>	1	10	10
14. section	6	21	11	34. <b>conferences</b>	2	10	9
15. couple	7	21	6	35. authors	2	10	7
16. <b>drafts</b>	8	20	9	36. creative	2	10	7
17. context	5	20	7	37. computers	2	8	7
18. final	7	20	7	38. <b>linked</b>	2	7	6
19. specific	5	19	8	39. minimum	2	7	6
20. function	1	18	18				

### Discussion

There were 39 AWL types that were repeated at a moderate or high level. These constitute the most likely candidates for incidental vocabulary acquisition from the teacher's speech. This seems to be a small amount, considering there are 3,110 types in the AWL. However, if students are exposed to this number of new

AWL types at the same rate throughout the term, it would result in the somewhat more substantial figure of 146 AWL types over the period of ten weeks. Considering the gains that students might have from attending to teacher speech in their other courses (such as reading or speaking/listening), it is possible that repetition-driven incidental acquisition results in students learning about as many as 10% of the AWL types in a fairly short period of time. On the other hand, there is strong evidence that the figure of 39 AWL types per eight classes should not be extrapolated.

The moderate and high repetition AWL types appear to be closely related to key terms in the course content or in common procedures, which are unlikely to change dramatically throughout the term. A majority of the types in table 4.5 clearly relate to either writing or grammar instruction (e.g., *thesis* and *paragraph*) or classroom procedures (e.g., *assignment* or *team*). These examples, like other high and moderate repetition AWL types, are inextricable from a grammar and writing course. It is not, therefore, safe to assume that students would be exposed to a new set of moderate and high repetition AWL types every two weeks.

This is further supported by the fact that many of these terms are used in multiple classes. For example, *tense* is used in six classes because the teacher

often discussed verb tense structure or referenced an assignment that pertained to it. This trend is not limited to a single instance: a majority of the AWL types in table 4.5 occur in four or more classes. For this reason, it is not likely that other class sessions in this course contain completely different high and moderate repetition of AWL types. Even if other class sessions contain a similar level of AWL use, they likely share some high-repetition items. Instead of extrapolating from this data, then, it is necessary to investigate the connection between classroom topics and AWL use. Instead of a focus on adding AWL types over time, I discuss the correlation between topic and classroom activities and repetition of AWL items.

It is worth pointing out the one high-repetition word that only occurs in one class. The type *function* is only used in one class, and is repeated heavily due to the teacher's presentation and discussion regarding content and function words. In fact, this term even receives an explicit vocabulary focus, as shown in the excerpt from the class it is used in (number seven):

alright the other kind of word are <EV> function  
 words <writes> the crucial difference between  
 content and function words are that for content  
 words the most important thing about them is what  
 they mean what's the definition for function  
 words do they have any meaning really what does  
 and mean could anyone answer that question

This shows that not all of the high-repetition AWL types occur in a well-distributed fashion, and that certain topics limited to one class session may create a relatively large amount of exposure to a given word. On the other hand, as stated above, explicit vocabulary focus does not play a large role in this course, nor is it associated with many AWL items.

It should be noted that this ability to compare AWL items across classes is an outcome of selecting contiguous classes. This focus on the development of the curriculum class by class makes a meaningful discussion of range possible. If I had chosen random samples throughout the course, it would have been difficult or impossible to find patterns in AWL use across classes.

Moving on from the high repetition category, the moderate repetition types seem linked to course content by the contexts of their use. A closer look at their context the teacher's speech reveals the connections to grammar, writing, and course procedures. For example, it seems plausible that *contrast* could arise in a number of different situations. The following excerpt is taken from class number seven, which featured ten repetitions of *contrast*, during the teacher's explanation of transition words:

what kind of transition I told some of you <ss>  
contrast it's a contrast similar to although  
 although although sname is brilliant he won't do  
 his homework and so he failed the class a



contrast to xx two two uh surprising or  
contrasted pieces of information

In this short excerpt, the students were exposed to three instances of *contrast*, one of *contrasted*, and two other AWL types. The sense of *contrast* in this example is explaining the function of a transition word, and then elaborating the meaning of an example sentence. So, even though there are many possible uses for *contrast*, its uses in these classes are related to explaining the function of contrasting information in writing. This tendency, for AWL items to be used in order to explain writing or grammar concepts, holds true for many of the other moderate repetition types.

In one case, a moderate repetition word with a limited range, *relevance*, shows that explicit vocabulary instruction may be correlated with repetition. *Relevance* occurs in only two classes, and is not a content-specific term in the same way that *editing* or *drafts* are, but it is tied to an explanation of classroom topics. Also, it is given an explicit vocabulary focus treatment, shown in the following excerpt from class:

does anyone remember what is a word that means about the same thing as relevance <ss> no that's that's one of the ways that you're going to do this but what's another word that has the same meaning as relevance <ss> importance importance right or in other words relevance is an academic word that means tell me so what who cares <ls> so we're back to that issue again so what who cares relevance is answering the question so what who

cares <1s> <writes> ok there are many ways to do  
 this there are I think three ways that are  
 appropriate to your level of writing

Here, there are four repetitions near the teacher's elicitation of a synonym for *relevance*, followed by a transition that links the term back to the focus of the teacher's lesson, ensuring that students focus on their topics and keep their essay writing relevant. In addition to being another example of the link between the repetition of AWL types and classroom topics, this example shows that explicit vocabulary focus may be associated with more repetition, even though this is one of the few AWL items that received this kind of treatment in class.

The lower-frequency and single categories of repetition have far more AWL types and contain the widest variety of AWL items, yet in terms of students' exposure to them through repetition in the teacher's speech, they are the least likely to be noticed and acquired. Furthermore, it is more difficult to characterize the low and single repetition AWL types because of their large number. However, they include more terms that are clearly related to the course content and procedures, such as *style*, *structure*, *partner*, *task*, and *schedule*. There are also more AWL types that are likely generally high-frequency in everyday spoken language, such as *finally* or *similar*, but even this pattern is hard to characterize, because there are fewer occurrences of these types in the teacher's talk.

Returning to the topic of which AWL items students may have learned about through incidental acquisition, the high and moderate repetition types also seem to be the clearest examples of words that students would have likely been familiar with at the beginning of class. The teacher does not explicitly teach or explain terms such as *draft* or *assignment*, but students show signs of understanding, such as following instructions and asking follow-up questions about assignments that show they understood the previous message. Ironically, then, the most frequently repeated AWL items are likely to be well understood by students. This point, established in the previous section, holds true for these most-repeated types. However, it could be that noticing these terms provides additional word knowledge, such as examples of contextual use.

Finally, it is interesting to note the differences in repetition between classes, even when accounting for variation in the amount of teacher talk, as shown by the normed rate in table 4.4. Given the seemingly strong connections between classroom topics and highly-repeated AWL types, one possible explanation is that the shifts in classroom topics and activities are responsible for the variation in AWL type repetition. It appears that understanding the relationship between classroom topics and AWL use may shed light into the

variation in the rate of repetition as well. This is one additional reason to turn towards the classroom topic category data in the next section.

#### **4.4 AWL Occurrence by Classroom Topics**

##### *Results*

The sections above have shown that there is a tendency for AWL types, especially high-repetition ones, to be connected with classroom content and procedures. Furthermore, the variation in the amount of teacher talk does not explain the difference in the rate of AWL use and repetition across classes. Some classes had a relatively large number of AWL types, despite having a smaller overall token count. In this section, I turn towards categories of classroom topics as a possible moderating influence on the exposure of students to AWL types.

Table 4.6 shows the number of tokens in each of the RANGE program's wordlists (i.e., 1,000 highest-frequency word families, second thousand highest-frequency word families, AWL families, and the remaining tokens that were not on any of the previous lists) for each of the activity codes. The row with AWL results is in bold. The raw count of AWL types ranges from zero (during teacher-dictated tests) to 666 (writing-focused language). The rankings of the wordlists are similar to the class-by-class results in table 4.1. The AWL has the fewest

tokens, while the one thousand wordlist has the most, followed by the two thousand word list, and off-list in decreasing order.

Table 4.6 shows that the number of AWL tokens is highest in the Writing-focused language category; it also has a high proportion of AWL tokens (3.5%). Although it features fewer tokens, the Pre-Written Text category also has a proportionally high occurrence of AWL tokens, accounting for 3.5% of the running words in the category (24 tokens). There were no AWL tokens in the Test category, and only 3 (1.7%) in the Anecdotes category, which are the two smallest categories in terms of total number of tokens. The next-smallest percentage of AWL tokens is found in the Activity Instructions category, which has only 118 AWL tokens (2.0% of the tokens in that category).

Table 4.6: Number of tokens per word list per classroom topic category									
List	Classroom Management	Activity Instructions	Assignments	Grammar Focus	Writing Focus	Pre-written Text	Test	Anecdotes	Total
#AWL	247 (2.3%)	118 (2%)	402 (3.3%)	174 (2.1%)	666 (3.5%)	24 (3.5%)	0 (0%)	3 (1.7%)	1634 (2.8%)
#1k List	9773 (91.1%)	5338 (91.8%)	10937 (88.9%)	7646 (90.7%)	17146 (89.8%)	581 (84.8%)	141 (89.2%)	157 (90.2%)	51719 (90.1%)
#2k List	414 (3.9%)	189 (3.2%)	527 (4.3%)	339 (4%)	686 (3.6%)	32 (4.7%)	13 (8.2%)	8 (4.6%)	2208 (3.8%)
#off-List	295 (2.7%)	172 (3%)	434 (3.5%)	275 (3.3%)	586 (3.1%)	48 (7%)	4 (2.5%)	6 (3.4%)	1820 (3.2%)
Total	10729	5817	12300	8434	19084	685	158	174	57381

Table 4.7 shows the number of AWL types and families by category, followed by the normed AWL types per hundred words. Writing-focused language and assignments feature the highest number of types (and families) with 189 (136) and 129 (95), respectively. The normalized count (AWL types / 100 words) shows that the number of types varies across categories relatively little, with the exception of pre-written texts, which is nearly three times the rate of the next-highest category (3.21 types / 100 tokens, compared to 1.15 types /100 tokens for anecdotes).

	<b>Classroom Management</b>	<b>Activity Instructions</b>	<b>Assignments</b>	<b>Grammar Focus</b>	<b>Writing Focus</b>	<b>Pre-written Text</b>	<b>Test</b>	<b>Anecdotes</b>	<b>Total</b>
<b>AWL Types</b>	98	56	129	71	189	22	0	2	334
<b>AWL Families</b>	75	42	95	61	136	20	0	2	213
<b>AWL Types / 100 tokens</b>	0.91	0.96	1.05	0.84	0.99	3.21	0.00	1.15	0.58

Finally, table 4.8 displays the ten most frequent AWL types in the categories with the largest number of AWL tokens: Classroom Management, Activity Instructions, Assignments, Grammar-focused language, and Writing-

focused language. Types that are in boldface do not occur in any of the other categories' top ten lists; for example, *conclusion*, *link*, *function*, *linking*, and *specific* only occur in the top ten writing-focused AWL types.

**Table 4.8: Most frequent AWL Types by classroom topic category  
(# of occurrences)**

<i>Classroom Management</i>	<i>Activity Instructions</i>	<i>Assignments</i>	<i>Grammar Focus</i>	<i>Writing Focus</i>
<b>assignments (15)</b>	thesis (14)	<b>assignment (33)</b>	passive (33)	paragraph (37)
couple (10)	journal (9)	passive (18)	<b>context (16)</b>	topic (34)
<b>conference (10)</b>	topic (7)	tense (16)	tense (10)	thesis (29)
drafts (9)	paragraph (6)	<b>section (15)</b>	<b>clause (8)</b>	transition (24)
<b>conferences (9)</b>	transition (5)	journal (14)	<b>plus (5)</b>	<b>conclusion (23)</b>
<b>draft (8)</b>	<b>topics (5)</b>	drafts (11)	<b>period (4)</b>	<b>link (20)</b>
thesis (7)	tense (4)	<b>project (11)</b>	team (4)	editing (18)
<b>grade (7)</b>	team (4)	final (10)	<b>structure (3)</b>	<b>function (18)</b>
final (6)	<b>identify (3)</b>	editing (10)	<b>job (3)</b>	<b>linking (17)</b>
editing (6)	couple (3)	<b>assigned (10)</b>	<b>similar (3)</b>	<b>specific (17)</b>

Types that do not occur in other category lists are in boldface.

### *Discussion*

The categories of classroom topics show interesting variation in terms of student exposure to AWL types. Some categories feature more exposure to AWL types than others, and the specific types that occur in each category tend to correspond to content and language functions related to that classroom topic. In this section, I interpret the classroom topic findings.

In terms of normed frequency of AWL types, the category of pre-written text features the most AWL tokens and types. This category stands apart from

the others because it is actually written text spoken aloud. It is not surprising that it has relatively more AWL use, because the AWL was created from a written corpus. Another factor could be that the teacher chose texts that challenge students with higher-level vocabulary. One example, from class seven, illustrates both these points; this pre-written text comes from a university textbook written for native English users:

whereas who cares literally asks you to identify a person or group who cares about your claims so what asks about the real world applications and consequences of those claims

This short excerpt contains three different AWL types, a fairly dense amount.

Earlier in the course, after assigning this text, the teacher asks how challenging students found the reading; most of them report that it was fairly difficult or very difficult for them. Thus, the high rate and variety of AWL types is not an accident, but linked to the lexical complexity of this text. Finally, it should be noted that this category does not account for a large amount of exposure, due to having a small total of tokens (685).

Writing-focused language features a large number of AWL tokens and types due to the number of text reference terms used in discussing process writing, and the presence of those same terms on the AWL. In the writing-focused portions of the classes, the teacher uses AWL types such as *paragraph*,



*topic, thesis, transition, and conclusion.* Because the AWL was compiled from written academic sources, it seems that references to textual features are well represented in the list, as well as in this category of classroom instruction.

Grammar-focused language does not include as much exposure to AWL items, but contains a set of AWL types that are strongly related to the topic of grammar. In grammar-focused parts of this class, students are exposed more often to terms such as *passive, context, tense, and clause.*

The assignments category has a surprisingly large number of AWL tokens and types, considering it corresponds to the teacher assigning and clarifying homework. However, this is not due to the level of detail or specificity in the teacher's instructions, but to the inclusion of descriptions of the assignments, which use AWL types that are used in referring to grammar and writing content. For example, the terms *passive* and *tense* occur in this category because the teacher referred to them in terms of specific assignments. This category also features one of the strongest AWL type collocations in the data, *final drafts*. These examples show that it is not necessarily the language of giving assignments and setting expectations that is associated with AWL exposure, but the fact that giving assignments required a recycling of terms from the course content.

The categories of classroom management and activity instructions do not feature as many AWL types and tokens as the assignments category, but they have the same pattern of exposure, where content words from grammar and writing work their way into these classroom topics. The main exception, for classroom management, is the inclusion of *conference* and *conferences*, as this category included the teacher's discussion of upcoming writing conferences. However, aside from this, these categories present a cross-section of content words from grammar and writing, and few (or no) distinctive patterns of AWL use of their own.

#### **4.5 Further discussion: AWL use in context**

The main focus of this thesis is on students' exposure to individual AWL types, but it also stands to reason that the context of AWL items is also important. For many AWL items, what defines their use in academic writing is their collocational pattern (i.e., what other words or syntactical patterns precede or follow them). From the previous sections, it is not clear whether the occurrences of AWL items in this classroom matched common collocations in academic writing. If they did, then this presents another opportunity for students to learn from exposure to teacher talk; if they did not, this suggests a limit of the usefulness of classroom input. In order to discuss this in more depth, I

investigated the collocational patterns of the most frequent AWL types in the main classroom activity categories (Classroom Management, Assignments, Activity Instructions, Grammar Focus, Writing Focus) and compared them with collocational patterns in academic writing found in the Corpus of Contemporary American English (COCA) (Davies, 2008-). What follows is a closer look at *assignment, thesis, passive, and paragraph*. In some cases, this EAP classroom has similar collocational patterns, but they are generally limited compared to the most frequent patterns in academic writing.

First, I look at *assignment* and *assignments*, which were the most common types in the classroom topic categories of Assignments and Classroom Management, respectively. These two types occur in the teacher's speech 70 times, and feature a fairly limited collocational pattern with only some *that*-clause post-modification. Regarding the use of *assignment/assignments* in the EAP classes:

- The most common left-collocate content words are *homework* and *grammar*
- Determiners were the most common pre-modifiers (e.g., *an, the, this*)
- The most common right-collocates include *and, but, so, and if*
- Four instances of *that*-clauses post-modifying *assignments* (e.g., "...assignments that you gave to me")

- Two instances of *for* (“...assignments for another class”).

Regarding the use of *assignment/assignments* in COCA:

- Also features *homework* as the most common left-collocate
- Other left-collocates include: *writing*, *group*, and *reading*
- However, COCA also includes the collocate *random*

Interestingly, the use of *assignment* to denote student homework occurs frequently in both this EAP course and in general academic writing. It seems likely that these students would also understand other common collocations from the COCA (e.g., *writing assignment*) because it uses *assignment* in the same sense, as homework. However, the phrase *random assignment*, from explanation of experimental designs, did not occur at all in the EAP course, where there was no occasion to discuss scientific methodology. This represents a very different meaning that students would be very unlikely to grasp.

In the category of activity instructions, *thesis* was the most common AWL type, occurring 59 times in the teacher’s speech throughout the transcribed classes. Regarding *thesis* in the EAP classes:

- The most common left-collocate is *your* (12 instances)
- There are only two right-collocates: *statement* and *statements*

The use of *your* reflects the spoken, direct nature of the classroom context.

The limited right-collocates show that this AWL item was used in a very narrow sense.

Regarding the use of *thesis* in COCA:

- The most frequent collocational patterns include *master's, doctoral, his, my,* and *M.A.*
- Other patterns that the EAP course did not feature include: *central, main, novel,* and *basic*
- COCA also shows the use of *thesis statement* as a frequent, notable pattern

Thus, students in the EAP class may have grasped one commonly used collocation, *thesis statement*. However, they were not exposed to other common patterns that refer to the central idea of an argument or a thesis as a document.

The most common AWL type in grammar-focused language was *passive*. All 53 instances refer to the grammatical construction of passive voice. This is supported by its common collocations in the teacher's speech. Regarding the use of *passive* in the EAP course:

- Common constructions include *passive voice* and *forming the passive*
- Every instance of *passive* is preceded by *a* or *the*

The teacher often said *the passive* as a short-hand reference to passive voice construction, which was one of the main grammar topics in the course. This pattern is also found in COCA's academic writing, but these instances are all in articles that discuss or give writing advice. Unsurprisingly, COCA shows that *passive* occurs in the sense of "not active" with a much wider range of nouns, including *smoking, recipients, learning, and resistance*.

The most common AWL type in writing-focused language was *paragraph*, which occurs 51 times in the teacher's speech throughout the classes. Common collocations in the EAP classes include:

- *body paragraph*
- *five paragraph essay*
- *determiner + paragraph*
- *effective paragraph*

COCA also has these collocations. However, it has one very frequent collocation that is not present in the teacher's speech: *in the \_\_\_\_\_ paragraph*, where the blank might be filled with *first, opening, preceding, previous, or final*. This use seems to serve the function of organizing and building cohesion in written text, and is thus not found commonly in the teacher's speech.

#### 4.6 Review of Findings

1. *How many AWL words are spoken by the teacher in one Grammar and Writing EAP classroom over two weeks of class?*

Students who attended these eight classes were exposed to a variety of AWL types (a total of 334 types and 213 families). The frequency of AWL use varied across classes (for a total of 1634 tokens). In general, exposure to AWL items is associated with the link between those types and the course's content and procedures. Many of these AWL types are common to English language courses, even at a beginning level, and thus are likely to be known by students at the beginning of the course, though it is possible that the use of these types provides incremental gains to their word knowledge. Finally, only a few instances of AWL exposure due to explicit vocabulary focus are in the data, suggesting that explicit vocabulary instruction is not the only activity that exposes students to high-value vocabulary in classrooms.

2. *How frequently are those AWL words are repeated in each class?*

The repetition of AWL types varies from class to class, but for the entire eight classes, 26 AWL types are repeated 6-10 times in at least one class, and 13 are repeated 11 or more times in at least one class. These moderate and high

repetition AWL types are generally related to core course topics (e.g., elements of essay writing and passive voice construction) or procedures (assigning homework, discussing the course schedule). Similar to research question one, the high repetition AWL types would likely be known to even beginning-level students familiar with English language instruction. Also, explicit vocabulary instruction may not account for many AWL types, but in at least one case, it is associated with a high repetition type, *function*.

3. *What classroom topics/routines feature the highest frequency of AWL tokens and the most AWL families?*

Writing-focused language has the largest number of AWL tokens (666) and types (189) due to the prevalence of textual reference terms used by the teacher during writing instruction. This reflects the intersection of the course content and the written-language origin of the AWL. Although pre-written text accounts for a small number of tokens, it contains a relatively dense concentration of AWL terms, perhaps due to this category's inclusion of written language and challenging readings. Classroom management, activity instructions, and assignments have a few distinct AWL types, but are fairly similar due to the inclusion of AWL types that relate to both writing and grammar instruction. In



addition, the classroom topic data helps confirm the connection between the occurrence of specific AWL types and course content and activities.

## **Chapter Five: Conclusion**

In this chapter, I present a summary of findings, implications for EAP teaching and learning, connections to the larger field of applied linguistics, the limitations of this thesis, and suggestions for future research. Finally, I conclude the chapter and this thesis.

### **5.1 Summary of Findings**

The previous chapter has shown that some AWL items are indeed part of classroom input that students are exposed to. The results show that they were not used randomly, but were connected to classroom topics and procedures. Specifically, AWL types in this teacher's speech were linked to content in writing and grammar instruction. This came through in part because the teacher used explicit instruction of topics. In terms of exposure to AWL items, then, it is clear that they are at least present in the teacher's speech, which means that students had at least a small chance of beginning to learn about a substantial number of AWL types. As Ellis (1994) notes, some word knowledge may be gained from hearing a word one time. It would be surprising if students learned a great deal about the AWL items spoken only once in these classes, but it is possible that students gained a base-level awareness of form or meaning of these terms.

The results show that a fairly limited set of AWL items are repeated enough for successful incidental acquisition to be likely. This suggests that by attending to the teacher's speech, students are not likely to learn a large portion of the AWL, but they will have a lot exposure to some items. Also, the teacher's speech featured the most repetition with terms that seem strongly connected to course objectives. There are many AWL items that were not repeated as much, likely because they do not have any clear connection to the topics of grammar, writing, or activity instruction. Because of lower repetition, AWL terms that are not connected to course content are unlikely to be learned well through incidental acquisition in courses.

In addition, there is evidence that students already understood most of the high and moderate level repetition AWL items before taking this course. These AWL items were used to explain activities, introduce content, and direct students. The students showed an understanding of the teacher's message by performing activities, discussing content in class, and responding to the teacher's commands. Furthermore, it seems likely that students who had previously studied English writing would be familiar with many of these terms, which include *paragraph*, *draft*, *conclusion*, and *assignment*. However, just because students understand a form-meaning connection does not mean that they have

mastered a word. It could be that the repetition of these items helped develop students' recognition of their spoken forms or other aspects of word knowledge.

The analysis of categories of classroom topics shows even more evidence that classroom content is an important influence on the AWL items that students are exposed to. For instance, the number of AWL tokens is greater in writing-focused language and pre-written text. So, it seems that students had greater exposure to AWL items because of the different types of classroom content, such as the combination of grammar and writing.

Even if students had the opportunity to hear AWL items repeatedly, notice them, and acquire some degree of word knowledge about them, the use of AWL items in this course did not always directly relate to the meanings or uses in academic writing. Any incidental learning of aural word form could still be helpful for students who go on to study academic language, but for some AWL types, students will have to learn additional meanings and constructions used in academic genres. For example, students who easily recognize the word *passive* because of its frequent use and repetition in this course would likely recognize the word in an academic context, yet without additional instruction or study, it is unlikely that they would make the semantic leap from "passive voice" to the meaning of "receptive, not taking action."

## 5.2 Implications for EAP teaching

The results of this thesis provide useful pedagogical implications for EAP education. Primarily, the use of AWL types in this study shows that students may gain some familiarity with AWL items by attending to their teacher's speech in the classroom. This familiarity could be useful in explicitly teaching AWL items. Teachers could try to activate this existing knowledge and build connections to it when they discuss the AWL. For example, when teachers address AWL (or other high-value, low-frequency) items explicitly, they should give students the opportunity to hear the spoken form of the items and the chance to think about whether the word sounds familiar. Even if students are not capable of recalling the meaning of the words accurately, this gives them the opportunity to access the low-level of knowledge they have about the word, which could help in learning new information, consolidating it, and recalling it successfully at a later time. Instead of just presenting students with word lists, this gives them the opportunity to recall any word knowledge they might have, thereby strengthening existing connections through the retrieval process.

In addition to asking students whether they have heard AWL words before, teachers should check to see if students are able to recall one (or more) of the word's meanings and use this as an anchor for new information. For example,

in the case of learning the most common senses of the term "passive" in academic prose, a teacher might start by asking students if they recognize this word from grammar class. Some AWL items will have closer meanings to those that students are familiar with, but many will have greatly divergent or contradictory meanings that will require instruction, translation, or dictionary definition. This kind of explicit focus on learning the AWL is important because it does not seem likely that students encounter all of the word list items through incidental exposure in their courses, based on the trends found in this thesis.

Also, teachers of AWL vocabulary courses could use the appendix from this thesis as a guide in identifying what AWL items students might have noticed from a grammar and writing course. Instead of poring over thousands of AWL types to try to guess or find assessment methods of what students might already know, teachers could review the appendix list of AWL types ranked by frequency in order to combine their intuition with concrete lexical usage patterns. Thus, the results of this thesis could be used as a tool to help raise teachers' awareness of what AWL items might be used in their institution. In turn, this could help them understand what students' needs are vis-a-vis AWL instruction.

The presence of AWL items in an intermediate-level course suggests that early on, students should have the tools to notice prominent words in spoken

language and incorporate them in their self-study of vocabulary. If students are able to notice prominent or frequently used words used by their teacher or other speakers and to make a guess at its written form, they may be able to make more use of the high-value vocabulary that occurs frequently in their classes or other sources of input. Thus, this supports giving students practice listening closely to lectures, practicing note-taking skills, and trying to spell unfamiliar words from aural input. Also, the earlier students become proficient in these skills, the more they might benefit from opportunities to learn AWL vocabulary through attending to their teacher's speech. This also applies to introducing students to vocabulary learning resources and strategies at an early stage in their studies. That is, students would benefit from learning dictionary skills and successful self-study strategies so they can make the most of the opportunities and rich input that they have.

The teacher of this course took many opportunities to explicitly focus on vocabulary that came up incidentally in the class, including several AWL items. This thesis has shown there may be a wealth of latent vocabulary practice in a grammar and writing course, and that these brief comprehension checks, elicitation of meaning, or presentations of word meaning may give students a chance to consolidate knowledge about words they have noticed through

repetition. In the results of this thesis, one AWL word fell into the high repetition category in part because the teacher used it several times in giving its definition and using it in examples. This shows that there is a value in using class time to focus on word meanings, even if there is not an explicit vocabulary learning component in assignments or assessments.

Given the link between classroom content and vocabulary, there may be a vocabulary-learning benefit for incorporating different topics in EAP courses and support for content-based language instruction. Even incorporating a series of thematic readings on a particular subject and discussing them in class may help enrich students' exposure to important vocabulary. As this thesis has shown, the content does not necessarily have to be academic or scientific in nature, nor does there have to be an explicit vocabulary focus in order to provide incidental vocabulary acquisition opportunities. At the same time, it is important to establish a degree of depth in each topic in order to provide opportunities for recycling of vocabulary. This thesis also provides additional rationale for EAP programs to have content-based language courses, because they may have substantial opportunities for incidental acquisition of academic vocabulary. If a grammar and writing course provides exposure to AWL items, it is likely that courses with academic content would also give students the opportunity to hear,



understand, and learn important academic vocabulary. Without a variety of content, students will not be exposed to as many AWL words.

Finally, it should be noted that this study should not be construed as support for pedagogical interventions designed to increase teachers' use of AWL words in the classroom. The AWL use that occurred in this course was unplanned and stemmed from core objectives and language functions. Any attempt to increase the amount of AWL use in a teacher's speech (e.g., through pre-written scripts to be read out loud) would not only miss the point of the findings of this thesis, but would likely fail.

### **5.3 Relation to other research**

This thesis contributes a detailed study of learners' exposure to high-value vocabulary spoken by a teacher in an English language classroom. A comparison to other research shows that this EAP classroom featured more high-value vocabulary items than some previously studied classrooms that were not EAP-focused. Table 5.1 provides a summary of these studies, the relevant measure of AWL use, and a comparison to the results of this thesis.

The clearest comparison is to Horst's 2009 study of a primary school EFL program in Quebec, wherein Horst reports the number of AWL families used in

each class. In that teaching context, each class featured a mean 27 AWL families used by the teacher, compared to a mean 66 per class in this thesis. At a minimum, this shows that AWL use is not a consistent level in all types of classrooms. It may be the case that features of the curriculum, such as the explicit teaching of process writing (as opposed to the communicative language teaching methods of the Quebec context), are responsible for the greater number of AWL

**Table 5.1: Comparison of AWL use in other classroom contexts**

Study	Context	# of types / tokens / % of AWL (or comparative)	Comparison to this thesis
Horst (2009)	Primary School ESL program in Quebec	Mean number of AWL families in teacher speech per class: <b>27</b>	Mean number of AWL families in teacher speech per class: <b>66</b>
Donzelli (2007)	Primary School EFL program in Italy for one year	AWL types as % of total types: <b>1.06%</b>	AWL types as % of total types: <b>12.99%</b>
Lightbown, Meara, & Halter (1998)	“Communicative classes of low intermediate adult learners with mixed first language backgrounds...” (p.224)	Mean number of UWL lemmas per 1500 word samples, ranging from 3.0 to 7.5 lemmas; off-list lemmas from 4.5 to 8.0 lemmas per 1500 word samples: <1.0% of tokens per sample	Percentage of AWL tokens per class: <b>1.8%-3.6%</b>
Dang & Webb (2013)	British Academic Spoken English (BASE) corpus	% AWL tokens: <b>3.82%</b> (Arts and Humanities genre) to <b>5.21%</b> (Social Sciences)	% AWL tokens: <b>2.8%</b>
Coxhead (2011) – review of multiple studies	Multiple studies investigating academic corpora	An average 10% coverage of academic texts.	% AWL tokens: <b>2.8%</b>

families in the IELP classroom. However, it may also be that divergent student needs in the two classrooms (adult EAP learners in one, Canadian primary school students in the other) account for the difference in AWL use by the teachers.

Another point of comparison comes from Donzelli's 2007 lexical analyses of a full year of a primary-level EFL course in Italy. In that study, Donzelli reported that AWL types comprised 1.06% of the number of types (note that Donzelli only reported the number of types, not tokens covered by the AWL). Although it is not reported above, the RANGE program automatically calculates this figure. The EAP teacher's speech featured 334 AWL types out of 2571 total types, or 12.99%. Again, the IELP classroom features more exposure to AWL items.

Finally, Lightbown, Meara, and Halter (1998) expanded their lexical profile study to include teacher talk in an adult education course that employed communicative methods. They found 3.0 to 7.5 lemmas from the University Word List per 1500 word samples from these courses. The UWL is not directly comparable to the AWL, but it does serve the same goals, of describing common vocabulary in academic contexts. Even if the off-list lemmas were added to Lightbown, Meara and Halter's figure in order to bring the UWL closer to the

makeup of the AWL, the result would be a range of 7.5 to 14.5 lemmas per 1500 words, or less than one percent of the running words. Again, compared to this, the EAP teacher's speech featured more AWL items, with a range of 1.8% to 3.6% of running words.

Aside from the quantitative comparisons, the results have shown that a fairly detailed focus and understanding of the setting is necessary in order to understand the role that incidental acquisition might play. Studies such as Tang (2011) have used very broad, coarse lexical measures to describe the degree of exposure to new or low-frequency words, which are unlikely to provide meaningful, useful answers. It would have been a mistake for me to simply report the number of AWL families used in these classes, extrapolate from that figure, and then make a judgment that they were either likely a good source of vocabulary learning or an inadequate one. This thesis has shown that the patterns of use and connections to classroom topics and functions can be crucial in interpreting findings, because, in this case, there were many AWL items that students would likely have known before taking this course.

Although the EAP course in this study featured more exposure to high-value vocabulary items than some previously studied classrooms, it did not have the same level of AWL use as mainstream academic lectures or in English

language instruction textbooks. In comparison with Dang and Webb's analysis of the British Academic Spoken English (BASE) corpus (2013), the EAP classes in this study featured fewer AWL tokens. Dang and Webb found that AWL use ranged from 3.82% (in the arts and humanities sub-corpus) to 5.21% (in the social sciences sub-corpus), with a figure of 4.41% coverage over the entire BASE corpus. The EAP teacher's talk may be simpler than university language in order to aid comprehension, or it may be that the topics and communicative functions in the EAP course are less tied to AWL items.

One potential argument against the importance of incidental vocabulary acquisition is that it is not possible if there is no new word knowledge present in the input learners are exposed to. This thesis provides a counter-example to that argument, as these classes featured hundreds of AWL types, and it is not likely that students were familiar with all of them. Also, I have made this case based on only the occurrence of AWL types in the class, unlike past research on classroom environments, which mostly focused on off-list types. If anything, the focus on AWL types has underestimated the opportunities for students to learn new words through incidental acquisition. Admittedly, this is not the strongest sort of evidence for the importance of incidental vocabulary acquisition, but it still adds

to the understanding of this construct's role in second language vocabulary acquisition.

#### **5.4 Limitations**

This section addresses the limits to this thesis' generalizability and its methodological limitations. I begin with the issue of sample size and generalizability. Secondly, I address the decision not to include student speech. Finally, I discuss the focus on spoken language.

Because of the specific focus of this thesis and the exploratory nature of the research, there are variables that I have not addressed, such as inter-teacher speaking style and course subject. The sample includes only one teacher, one classroom, one subject, and one curriculum. This sample clearly does not offer a statistically generalizable account of the frequency of AWL use in EAP programs. Though this focus means that it would be improper to draw global conclusions about EAP classrooms in general, it also means that this thesis provides a detailed account of AWL use over eight classes and thorough context about the course. Teachers and researchers may recognize similarities between this classroom and other EAP contexts, and may find that this thesis provides a useful transferrable knowledge about incidental AWL item acquisition.

In addition to the focus on one classroom, the data in this study are limited to teacher talk, and exclude student speech. This is a study of AWL use in teacher talk, and does not directly include students' speech, even though it is possible that attending to other students' speech may be another source of incidental acquisition of vocabulary. For this reason, it should be noted that the quantitative results may actually under-count AWL learning opportunities, because they may occur in student speech as well.

Finally, this study focused only on receptive oral language, and excluded other opportunities for vocabulary acquisition. It is clear from the video recordings that students were exposed to AWL items that occurred in written text on the whiteboard, handouts, textbooks, and student writing. A study that integrated these sources could paint a much more complex picture of vocabulary acquisition, incorporating different patterns of support and repetition across language modes.

### **5.5 Future research**

There are several avenues for future research projects related to this thesis. As mentioned in the section on limitations, this thesis provided a detailed description of one teacher, classroom, curriculum, and program. The evidence for incidental acquisition of AWL items in EAP classrooms could be refined or

strengthened by studies of other teachers, classrooms, and curricula, or by studies that investigate several classrooms simultaneously, thereby accounting for more variation.

In addition to investigating other classroom contexts, future research could look at a wider range of students' exposure, incorporating speaking, listening, reading, and writing. There is a need for studies to investigate connections between students' exposure to spoken language and written language in various contexts in order to better understand how much new vocabulary EAP students are exposed to, and the role of repetition and recycling across modes. This could also include students interacting with other students, a source of language exposure that this study did not include.

Even though EAP programs help students prepare for the challenges of using English in academic environments, there is scarce information about vocabulary that is common in spoken academic English. Some research has reported on the use of the AWL in academic speech, yet the AWL is an inappropriate gauge of spoken vocabulary because it was derived from written language. If a spoken academic word list was compiled, perhaps using corpora such as the BASE or MICASE, researchers could investigate the similarities and differences between EAP programs and various academic environments, and ask



whether students in EAP programs are exposed to similar types of spoken input. Such a study would require a first step of compiling larger, more diverse spoken academic corpora.

Finally, this thesis investigated the AWL items that students were exposed to in one EAP classroom, but did not measure students' vocabulary knowledge. Being able to have a combination of analyzing students' exposure to AWL items (or other high-value vocabulary) and measuring the changes in their word knowledge over time would provide a test of the strength of incidental vocabulary acquisition from spoken language.

## **5.6 Conclusion**

This study examined students' exposure to AWL items through their teacher's speech in one curricular unit of an EAP intermediate grammar and writing course in order to better understand which high-value vocabulary students might acquire through attending to the teacher and noticing words that are used, in particular those words repeated often. Through this focused account of one EAP classroom, I have shown that there were connections and trends between topics in class, classroom procedures, use of AWL items, and repetition of AWL items. Based on the quantitative results, it is indeed possible that

students in this EAP classroom made incremental gains in their knowledge of AWL items. However, it is unlikely that they learned a great deal about AWL items that they were not already familiar with.

With these results, I hope others interested in second language vocabulary acquisition in EAP programs will be able to better understand similar classroom environments and their potential for incidental vocabulary acquisition. Most importantly, it is my hope that understanding this issue better may help advance the quality of instruction and student success in EAP programs.

### References

- Bonk, W. J. (2000). Second language lexical knowledge and listening comprehension. *International Journal of Listening*, 14(1), 14–31.
- Brezina, V., & Gablasova, D. (2013). Is there a core general vocabulary? Introducing the new General Service List. *Applied Linguistics*, Advance Access published August 26, 2013. doi:10.1093/applin/amt018
- Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, 34(2), 213–238.
- Coxhead, A. (2011). The Academic Word List 10 Years On: Research and Teaching Implications. *TESOL Quarterly*, 45(2), 355–362.
- Dang, T. N. Y., & Webb, S. (2013). The lexical profile of academic spoken English. *English for Specific Purposes*. <http://dx.doi.org/10.1016/j.esp.2013.08.001>
- Davies, Mark. (2008-) *The Corpus of Contemporary American English: 450 million words, 1990-present*. Available online at <http://corpus.byu.edu/coca/>
- Donzelli, G. (2007). Foreign language learners: Words they hear and words they learn: A case study. *Estudios de Linguística Inglesa Aplicada (ELIA)*, 7 (1), 103–125.
- Ellis, R. (1994). Factors in the Incidental Acquisition of Second Language Vocabulary from Oral Input: A Review Essay. *Applied Language Learning*, 5(1), 1–32.
- Ellis, R., Tanaka, Y., & Yamazaki, A. (1994). Classroom interaction, comprehension, and the acquisition of L2 word meanings. *Language Learning*, 44(3), 449.
- Elisha-Primo, I., Sandler, S., Goldfrad, K., Ferenz, O., & Perpignan, H. (2010). Listening to students' voices: A curriculum renewal project for an EFL graduate academic program. *System*, 38(3), 457–466.
- Ferris, D., & Tagg, T. (1996). Academic listening/speaking tasks for ESL students: Problems, suggestions, and implications. *TESOL Quarterly*, 30(2), 297–320.

- File, K. A., & Adams, R. (2010). Should Vocabulary Instruction Be Integrated or Isolated? *TESOL Quarterly*, 44(2), 222–249.
- Folse, K. (2010). Is explicit vocabulary focus the reading teacher's job? *Reading in a Foreign Language*, 22(1), 139-160.
- Gardner, D., & Davies, M. (2013). A new academic vocabulary list. *Applied Linguistics*, Advance Access published August 2, 2013.  
doi:10.1093/applin/amt015
- Goh, C. (1999). How much do learners know about the factors that influence their listening comprehension? *Hong Kong Journal of Applied Linguistics*, 4(1), 17–41.
- Heatley, A., Nation, I.S.P. & Coxhead, A. (2002). *RANGE program*. Downloaded from: [http://www.vuw.ac.nz/lals/staff/Paul\\_Nation](http://www.vuw.ac.nz/lals/staff/Paul_Nation)
- Horst, M. (2009). Revisiting classrooms as lexical environments. In Fitzpatrick, T. & Barfield, A. (Eds.), *Lexical Processing in Second Language Learners: Papers and Perspectives in Honour of Paul Meara* (pp. 53–66). Bristol, UK; Buffalo: Multilingual Matters.
- Horst, M. (2010). How well does teacher talk support incidental vocabulary acquisition? *Reading in a Foreign Language*, 22(1). Retrieved from <http://spectrum.library.concordia.ca/36106/>
- Hsu, J. T. (2010). The effects of collocation instruction on the reading comprehension and vocabulary learning of college English majors. *Asian EFL Journal*, 12(1), 47–87.
- Hsueh-Chao, M. H., & Nation, P. (2000). Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language*, 13(1), 403–30.
- Hulstijn, J. H. (1992). Retention of inferred and given word meanings: Experiments in incidental learning. In Arnaud, P.J.L. & Bejoint, H. (Eds.), *Vocabulary and Applied Linguistics* (pp. 113-125). London: Macmillan.

- Hulstijn, J. H. (2001). Intentional and incidental second language vocabulary learning: A reappraisal of elaboration, rehearsal and automaticity. In P. Robinson (Ed.), *Cognition and Second Language Instruction* (pp. 258-286). Cambridge: Cambridge University Press.
- Hyland, K., & Tse, P. (2007). Is there an “academic vocabulary”? *TESOL Quarterly*, 41(2), 235–253.
- Joe, A. (2010). The quality and frequency of encounters with vocabulary in an English for academic purposes programme. *Reading in a Foreign Language*, 22(1), 117–138.
- Keating, G. D. (2008). Task effectiveness and word learning in a second language: The involvement load hypothesis on trial. *Language Teaching Research*, 12(3), 365–386.
- Krashen, S. (1982). *Principles and practice in second language acquisition*. Pergamon Press: Elmsford, NY.
- Laufer, B. (1989). What percentage of text lexis is essential for comprehension? In C. Lauren & M. Nordman (Eds.), *Special Language: From Humans Thinking to Thinking Machines* (pp. 316–323). Clevedon, U.K.: Multilingual Matters.
- Laufer, B., & Hulstijn, J. H. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics*, 22(1), 1–26.
- Laufer, B., & Ravenhorst-Kalovski, G. C. (2010). Lexical threshold revisited: Lexical text coverage, learners’ vocabulary size and reading comprehension. *Reading in a Foreign Language*, 22(1), 15–30.
- Lightbown, P. M., Meara, P., & Halter, R. (1998). Contrasting patterns in classroom lexical environments. In Albrechtsen, D., Henriksen, B., Poulsen, E., & Mees, I. (Eds.), *Perspectives on foreign and second language pedagogy* (pp. 221–238). Odense, Denmark: Odense University Press.
- Long, M. H. (1981). Input, interaction, and second-language acquisition. *Annals of the New York Academy of Sciences*, 379(1), 259–278.

- Lynch, T. (2011). Academic listening in the 21st century: Reviewing a decade of research. *Journal of English for Academic Purposes*, 10(2), 79–88.
- Meara, P., Lightbown, P. M., & Halter, R. H. (1997). Classrooms as lexical environments. *Language Teaching Research*, 1(1), 28–46.
- Miller, D. (2011). ESL reading textbooks vs. university textbooks: Are we giving our students the input they may need? *Journal of English for Academic Purposes*, 10(1), 32–46.
- Milton, J. (2009). *Measuring second language vocabulary acquisition*. New York: Cambridge University Press.
- Mozejko, Z.P. (2011). Student needs assessment in teaching English at the tertiary level: An individual learner differences perspective. In Arabski, J. & Wojtaszek, A. (Eds.), *Individual Learner Differences in SLA*, (pp. 184-198). Clevedon, U.K.: Multilingual Matters.
- Nagy, W. E., Herman, P. A., & Anderson, R. C. (1985). Learning Words from Context. *Reading Research Quarterly*, 20(2), 233–253.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. New York: Cambridge University Press.
- Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review/La revue canadienne des langues vivantes*, 63(1), 59–82.
- Portland State University Intensive English Language Program (2012). *IELP course aims by level*. Portland, Oregon.
- Reder, S., Harris, K., & Setzler, K. (2003). The multimedia adult ESL learner corpus. *TESOL Quarterly*, 37(3), 546–557.
- Rost, M. (1994). On-line summaries as representations of lecture understanding. In Flowerdew, J. (Ed.), *Academic listening: Research perspectives* (pp. 93-127). New York: Cambridge University Press.

- Schmitt, D. and Schmitt, N. (2005). *Focus on vocabulary: Mastering the Academic Word List*. White Plains, NY: Longman Press.
- Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. New York: Palgrave Macmillan.
- Schmitt, N., & Meara, P. (1997). Researching vocabulary through a word knowledge framework. *Studies in Second Language Acquisition*, 19(1), 17–36.
- Schmitt, N., Jiang, X., & Grabe, W. (2011). The percentage of words known in a text and reading comprehension. *The Modern Language Journal*, 95(1), 26–43.
- Staehr, L. S. (2008). Vocabulary size and the skills of listening, reading and writing. *Language Learning Journal*, 36(2), 139–152.
- Spada, N., & Fröhlich, M. (1995). *COLT: Communicative orientation of language teaching observation scheme coding conventions and applications*. National Centre for English Language Teaching and Research, Macquarie University.
- Tang, E. (2011). Non-native teacher talk as lexical input in the foreign language classroom. *Journal of Language Teaching and Research*, 2(1), 45–54.
- Texas A&M University English Language Institute (2007). *Course descriptions*. Retrieved from <http://eli.tamu.edu/program/courses.php>
- University of Lincoln English Language Centre (2013). *Pre-sessional English and academic study (PEASS) courses*. Retrieved from <http://www.lincoln.ac.uk/home/international/englishlanguagecentre/peass>
- University of Pennsylvania English Language Programs (2013). *800 level courses*. Retrieved from [http://www.sas.upenn.edu/elp/intensive\\_advanced\\_800](http://www.sas.upenn.edu/elp/intensive_advanced_800)
- Vidal, K. (2003). Academic listening: A source of vocabulary acquisition? *Applied Linguistics*, 24(1), 56–89.

- Webb, S. (2007). The effects of repetition on vocabulary knowledge. *Applied Linguistics*, 28(1), 46–65.
- Webb, S., & Kagimoto, E. (2009). The effects of vocabulary learning on collocation and meaning. *TESOL Quarterly*, 43(1), 55–77.
- Webb, S., & Rodgers, M. P. H. (2009). Vocabulary demands of television programs. *Language Learning*, 59(2), 335–366. doi:10.1111/j.1467-9922.2009.00509.x
- West, M. (1953). *A general service list of English words: With semantic frequencies and a supplementary word-list for the writing of popular science and technology*. London: Longman, Green, & Company.
- Xue, G., & Nation, I. S. P. (1984). A university word list. *Language learning and communication*, 3(2), 215–229.
- Zahar, R., Cobb, T., & Spada, N. (2001). Acquiring vocabulary through reading: Effects of frequency and contextual richness. *Canadian Modern Language Review/La Revue canadienne des langues vivantes*, 57(4), 541–572.



## Appendix A: Institutional Review Board Approval Letter



Post Office Box 751  
Portland, Oregon 97207-0751  
Human Subjects Research Review Committee  
hsrrc@lists.pdx.edu

503-725-2227 tel  
503-725-8170 fax

Date: July 1, 2013

To: Eric Dodson

From: Todd Bodner, HSRRC Chair

Re: HSRRC approval for your project titled, "Opportunities for Incidental Vocabulary Acquisition in an English for Academic Purposes Classroom" (HSRRC Proposal # 132669)

Review Type: Exempt, Category 4

Dear Eric,

In accordance with your request, the Human Subjects Research Review Committee has reviewed your proposal referenced above for compliance with PSU and DHHS policies and regulations covering the protection of human subjects. The Committee is satisfied that your provisions for protecting the rights and welfare of all subjects participating in the research are adequate, and your project is approved. Please note the following requirements:

**Changes to Protocol:** Any changes in the proposed study, whether to procedures, survey instruments, consent forms or cover letters, must be outlined and submitted to the Committee immediately. The proposed changes cannot be implemented before they have been reviewed and approved by the Committee.

**Adverse Reactions and/or Unanticipated Problems:** If any adverse reactions or unanticipated problems occur as a result of this study, you are required to notify the Committee immediately. If the issue is serious, approval may be withdrawn pending an investigation by the Committee.

**Completion of Study:** Please notify the Committee as soon as your research has been completed. Study records, including protocols and signed consent forms for each participant, must be kept by the investigator in a secure location for three years following completion of the study (or per any requirements specified by the project's funding agency).

If you have questions or concerns, please contact the HSRRC in the Office of Research and Strategic Partnerships (RSP) at 503-725-2227, Market Center Building, Suite 620.

Cc: Susan Conrad, Advisor

### Appendix B: AWL Type List from All Classes

The types have been sorted by their frequency across all classes.

Type	FREQ	RANGE	Max. Freq.in One Class	1	2	3	4	5	6	7	8
thesis	59	7	17	2	13	17	17	1	7	2	0
passive	53	6	32	4	0	3	1	0	3	10	32
paragraph	52	6	18	5	1	4	13	0	18	11	0
topic	51	6	20	3	20	3	12	0	7	6	0
assignment	46	7	16	1	6	14	16	0	5	3	1
editing	36	7	8	6	1	8	0	8	8	3	2
tense	35	6	13	5	4	13	4	0	8	0	1
transition	30	3	25	0	2	0	3	0	0	25	0
journal	28	6	11	3	0	3	3	4	4	0	11
conclusion	27	3	23	0	2	0	0	0	23	2	0
link	25	4	14	2	2	0	7	0	0	14	0
assignments	23	5	7	0	5	7	7	2	2	0	0
linking	21	3	17	1	0	0	3	0	0	17	0
section	21	6	11	2	0	11	5	0	1	1	1
couple	21	7	6	3	0	5	2	1	3	1	6
drafts	20	8	9	1	1	1	1	2	1	4	9
context	20	5	7	7	2	0	3	0	5	3	0
final	20	7	7	3	1	0	1	1	2	5	7
specific	19	5	8	3	0	8	5	0	2	1	0
function	18	1	18	0	0	0	0	0	0	18	0
contrast	18	4	10	1	0	0	6	0	1	10	0
draft	18	6	8	1	0	0	1	3	1	4	8
process	17	6	6	6	2	1	0	0	5	1	2
paragraphs	17	6	4	3	4	1	4	0	3	2	0
job	16	5	6	6	0	3	1	0	4	2	0
assigned	16	5	4	0	3	4	3	2	0	0	4
relevance	15	2	9	0	0	0	0	0	6	9	0
topics	15	4	8	3	8	1	3	0	0	0	0
team	14	4	6	4	6	0	1	0	3	0	0
grade	14	6	5	2	0	5	0	1	1	1	4
academic	14	5	4	1	4	0	4	0	2	3	0

Type	FREQ	RANGE	Max. Freq.in One Class	1	2	3	4	5	6	7	8
structure	14	6	4	2	3	4	3	0	1	1	0
project	13	3	6	0	0	6	3	0	4	0	0
edit	12	4	9	1	0	0	1	1	9	0	0
parallel	12	2	8	0	0	8	4	0	0	0	0
task	12	5	4	0	0	2	2	4	0	1	3
conclusions	11	2	10	0	0	0	0	0	10	1	0
clause	11	3	8	0	2	1	0	0	0	8	0
conference	10	1	10	0	0	0	0	0	0	0	10
conferences	10	2	9	0	0	0	0	0	0	1	9
authors	10	2	7	7	0	0	0	0	0	3	0
creative	10	2	7	0	0	7	0	0	3	0	0
schedule	10	5	5	0	0	2	5	1	0	1	1
period	10	5	4	3	1	0	0	0	1	4	1
plus	10	5	4	4	1	2	0	1	0	2	0
identify	9	5	4	1	2	1	0	0	0	4	1
partner	9	4	3	2	0	0	0	3	1	3	0
similar	9	5	3	3	1	1	1	0	0	3	0
style	9	5	3	3	1	1	0	0	2	2	0
computers	8	2	7	1	0	0	7	0	0	0	0
concept	8	3	5	1	0	0	2	0	0	5	0
chapter	8	4	4	0	1	2	0	0	0	4	1
finally	8	6	3	1	1	1	1	0	3	0	1
journals	8	5	3	0	0	1	2	3	1	0	1
obvious	8	5	3	2	1	0	3	0	1	0	1
individual	8	5	2	0	2	1	2	0	2	0	1
linked	7	2	6	1	0	0	0	0	0	6	0
minimum	7	2	6	0	0	0	1	0	6	0	0
culture	7	3	5	1	1	0	0	0	0	5	0
chart	7	3	3	0	0	3	1	0	0	3	0
comments	7	5	2	2	0	2	0	1	1	1	0
instructions	7	5	2	2	1	1	0	0	1	0	2
comment	6	2	5	5	0	0	1	0	0	0	0
definition	6	2	5	1	0	0	0	0	0	5	0
computer	6	3	4	1	0	4	0	0	1	0	0
links	6	3	4	1	0	1	0	0	0	4	0
assume	6	2	3	0	0	0	3	0	0	3	0

Type	FREQ	RANGE	Max. Freq.in One Class	1	2	3	4	5	6	7	8
normally	6	4	3	0	0	3	1	0	0	1	1
tasks	6	3	3	0	0	3	0	1	0	0	2
individuals	5	1	5	0	0	0	5	0	0	0	0
significance	5	1	5	0	0	0	0	0	0	5	0
sections	5	2	4	0	0	4	1	0	0	0	0
transitions	5	2	4	0	4	0	0	0	0	1	0
author	5	2	3	0	0	0	0	0	2	3	0
credit	5	2	3	0	0	2	0	0	3	0	0
focus	5	3	3	1	0	1	0	0	0	3	0
goal	5	3	3	0	0	3	0	0	0	1	1
positive	5	3	3	3	0	0	0	0	1	1	0
strategies	5	3	3	1	0	0	0	0	1	3	0
whereas	5	3	3	1	0	0	0	0	1	3	0
appropriate	5	4	2	1	1	0	0	0	2	0	1
clauses	5	4	2	0	1	1	0	0	1	2	0
specifically	5	3	2	2	0	0	0	0	1	2	0
approach	4	1	4	0	0	0	4	0	0	0	0
duration	4	1	4	4	0	0	0	0	0	0	0
prediction	4	1	4	0	0	0	0	0	4	0	0
widespread	4	1	4	0	0	0	4	0	0	0	0
define	4	2	3	0	0	3	0	0	0	1	0
issue	4	2	3	0	0	0	1	0	3	0	0
jobs	4	2	3	3	0	0	0	0	0	1	0
major	4	2	3	3	0	1	0	0	0	0	0
odd	4	2	3	0	0	0	0	0	1	3	0
percent	4	2	3	0	0	3	1	0	0	0	0
accurate	4	3	2	1	0	1	0	0	2	0	0
designed	4	3	2	1	0	0	0	0	0	1	2
items	4	4	1	1	1	1	0	1	0	0	0
concluding	3	1	3	0	0	0	0	0	3	0	0
economic	3	1	3	3	0	0	0	0	0	0	0
environment	3	1	3	0	0	0	3	0	0	0	0
maximum	3	1	3	0	0	0	0	0	0	0	3
published	3	1	3	3	0	0	0	0	0	0	0
restricted	3	1	3	3	0	0	0	0	0	0	0
alternatively	3	2	2	0	0	2	0	0	0	0	1



Type	FREQ	RANGE	Max. Freq.in One Class	1	2	3	4	5	6	7	8
clarifying	2	1	2	0	0	0	0	0	0	2	0
definitions	2	1	2	0	0	0	0	0	0	2	0
dramatic	2	1	2	0	0	0	2	0	0	0	0
editor	2	1	2	2	0	0	0	0	0	0	0
emphasize	2	1	2	0	0	0	0	0	2	0	0
establish	2	1	2	0	0	0	0	0	0	2	0
evidence	2	1	2	0	0	2	0	0	0	0	0
expanded	2	1	2	0	0	0	0	0	2	0	0
focused	2	1	2	0	0	0	0	0	0	2	0
focuses	2	1	2	0	0	0	0	0	0	2	0
format	2	1	2	0	0	2	0	0	0	0	0
functions	2	1	2	2	0	0	0	0	0	0	0
gender	2	1	2	0	0	0	2	0	0	0	0
ignore	2	1	2	0	2	0	0	0	0	0	0
immigrants	2	1	2	2	0	0	0	0	0	0	0
immigration	2	1	2	2	0	0	0	0	0	0	0
label	2	1	2	2	0	0	0	0	0	0	0
margin	2	1	2	0	0	0	0	0	0	2	0
military	2	1	2	2	0	0	0	0	0	0	0
negative	2	1	2	2	0	0	0	0	0	0	0
occupations	2	1	2	0	0	0	0	0	0	2	0
perceptions	2	1	2	0	0	0	2	0	0	0	0
requirements	2	1	2	0	0	0	2	0	0	0	0
roles	2	1	2	0	0	0	2	0	0	0	0
teams	2	1	2	0	0	0	2	0	0	0	0
technology	2	1	2	0	0	0	2	0	0	0	0
text	2	1	2	0	0	0	0	0	0	2	0
tradition	2	1	2	0	2	0	0	0	0	0	0
affect	2	2	1	1	0	0	0	0	1	0	0
analysis	2	2	1	0	0	0	1	1	0	0	0
appreciate	2	2	1	0	0	0	0	0	0	1	1
assumed	2	2	1	0	0	0	1	0	0	0	1
aware	2	2	1	1	1	0	0	0	0	0	0
benefit	2	2	1	1	0	0	1	0	0	0	0
conflict	2	2	1	0	0	0	0	0	0	1	1
contact	2	2	1	0	1	0	0	0	0	1	0



Type	FREQ	RANGE	Max. Freq.in One Class	1	2	3	4	5	6	7	8
approaches	1	1	1	0	0	0	1	0	0	0	0
approaching	1	1	1	0	0	0	1	0	0	0	0
approximately	1	1	1	0	0	0	1	0	0	0	0
area	1	1	1	0	0	0	0	0	0	0	1
areas	1	1	1	0	0	1	0	0	0	0	0
automatically	1	1	1	1	0	0	0	0	0	0	0
beneficial	1	1	1	0	0	0	0	0	1	0	0
brief	1	1	1	0	0	0	0	0	0	1	0
categories	1	1	1	0	0	1	0	0	0	0	0
challenge	1	1	1	0	0	0	1	0	0	0	0
challenging	1	1	1	0	1	0	0	0	0	0	0
chapters	1	1	1	0	1	0	0	0	0	0	0
civil	1	1	1	0	0	0	0	0	0	1	0
commenting	1	1	1	0	0	0	1	0	0	0	0
commitments	1	1	1	0	0	0	0	0	0	0	1
community	1	1	1	0	1	0	0	0	0	0	0
compound	1	1	1	0	0	1	0	0	0	0	0
compute	1	1	1	0	0	0	0	0	0	0	1
conceptually	1	1	1	1	0	0	0	0	0	0	0
concludes	1	1	1	0	0	0	0	0	1	0	0
confined	1	1	1	0	0	0	0	0	0	1	0
confirm	1	1	1	0	0	0	0	0	0	1	0
construction	1	1	1	0	0	0	1	0	0	0	0
constructive	1	1	1	0	0	0	0	0	0	1	0
contrasted	1	1	1	0	0	0	0	0	0	1	0
conventions	1	1	1	1	0	0	0	0	0	0	0
convincing	1	1	1	0	1	0	0	0	0	0	0
correspondence	1	1	1	0	0	0	0	0	0	1	0
cultural	1	1	1	0	0	0	0	0	0	1	0
cycle	1	1	1	1	0	0	0	0	0	0	0
demonstrative	1	1	1	0	0	0	0	0	0	1	0
detective	1	1	1	0	0	0	0	0	0	0	1
device	1	1	1	0	0	0	0	0	0	1	0
distinction	1	1	1	1	0	0	0	0	0	0	0
distinctions	1	1	1	1	0	0	0	0	0	0	0
drafting	1	1	1	0	1	0	0	0	0	0	0





Type	FREQ	RANGE	Max. Freq.in One Class	1	2	3	4	5	6	7	8
item	1	1	1	0	0	0	1	0	0	0	0
labels	1	1	1	0	0	1	0	0	0	0	0
located	1	1	1	0	0	1	0	0	0	0	0
methods	1	1	1	1	0	0	0	0	0	0	0
modifying	1	1	1	0	1	0	0	0	0	0	0
nevertheless	1	1	1	0	0	0	0	0	0	1	0
obviously	1	1	1	0	0	1	0	0	0	0	0
occur	1	1	1	0	0	0	1	0	0	0	0
ongoing	1	1	1	0	0	0	1	0	0	0	0
option	1	1	1	0	0	1	0	0	0	0	0
optional	1	1	1	0	0	0	1	0	0	0	0
overall	1	1	1	0	0	1	0	0	0	0	0
paraphrasing	1	1	1	1	0	0	0	0	0	0	0
partners	1	1	1	0	0	0	0	0	0	1	0
precisely	1	1	1	0	0	0	0	0	0	1	0
previous	1	1	1	0	0	0	0	0	0	1	0
principle	1	1	1	0	0	0	1	0	0	0	0
prioritize	1	1	1	0	0	1	0	0	0	0	0
professional	1	1	1	0	0	1	0	0	0	0	0
proportion	1	1	1	0	0	0	1	0	0	0	0
random	1	1	1	0	0	0	0	0	0	1	0
refine	1	1	1	0	0	0	0	0	0	1	0
refined	1	1	1	0	0	0	0	0	0	1	0
rely	1	1	1	0	0	0	0	0	1	0	0
restriction	1	1	1	0	0	0	0	0	0	0	1
reverse	1	1	1	0	0	0	0	0	1	0	0
revising	1	1	1	0	0	0	0	0	0	0	1
series	1	1	1	0	0	0	0	0	0	1	0
significantly	1	1	1	0	0	0	0	0	1	0	0
similarly	1	1	1	0	1	0	0	0	0	0	0
specified	1	1	1	1	0	0	0	0	0	0	0
styles	1	1	1	0	0	0	1	0	0	0	0
stylish	1	1	1	0	0	0	0	0	0	1	0
subordination	1	1	1	0	0	0	0	0	0	1	0
subsequent	1	1	1	0	0	0	0	0	0	1	0
sufficient	1	1	1	0	0	0	1	0	0	0	0

Type	FREQ	RANGE	Max. Freq.in One Class	1	2	3	4	5	6	7	8
symbol	1	1	1	0	0	0	0	0	1	0	0
tape	1	1	1	0	0	1	0	0	0	0	0
taped	1	1	1	0	0	1	0	0	0	0	0
targeted	1	1	1	0	0	1	0	0	0	0	0
technological	1	1	1	0	0	0	1	0	0	0	0
tension	1	1	1	0	0	0	0	0	0	1	0
texts	1	1	1	1	0	0	0	0	0	0	0
transportation	1	1	1	0	0	0	1	0	0	0	0
underestimate	1	1	1	0	1	0	0	0	0	0	0
unspecified	1	1	1	1	0	0	0	0	0	0	0
variations	1	1	1	0	0	0	0	0	0	1	0
visualizing	1	1	1	1	0	0	0	0	0	0	0
visually	1	1	1	1	0	0	0	0	0	0	0
volunteer	1	1	1	0	0	0	1	0	0	0	0
volunteers	1	1	1	0	0	0	1	0	0	0	0