Spring 5-19-2014

A Comparison of Linguistic Features in the Academic Writing of Advanced English Language Learner and English First Language University Students

Margo K. Russell
Portland State University

A Comparison of Linguistic Features in the Academic Writing of Advanced English

Language Learner and English First Language University Students

by

Margo K. Russell

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
Lynn Santelmann
Linnea Spitzer

Portland State University
2014

Abstract

Writing for an academic purpose is not an easy skill to master, whether for a native English speaker (L1) or an English language learner (ELL). In order to better prepare ELL students for success in mainstream content courses at the university level, more must be known about the characteristics of student writing in the local context of an intensive English program. This information can be used to inform ELL writing instructors of which linguistic features to target so that their students produce writing that sounds appropriate for the academic written register.

Two corpora of 30 research essays each were compiled, one of L1 student writing done in various departments at Portland State University, and the other of ELL writing produced in an advanced writing course in Portland State University's Intensive English Language Program. The corpora were compared for the frequencies of 13 linguistic features which had been previously found in significantly different frequencies in L1 and ELL essays (Hinkel, 2002).

The tokens of each feature in each essay were counted, and the frequency rate was calculated in each case. The results of the Mann-Whitney U test found 6 features with significantly different frequency rates between the two corpora. The following features were more frequent in L1 essays than in ELL essays: modal *would,* perfect aspect, passive voice, reduced adjective clause, and *it*-cleft. In addition, the type/token ratio was found be significantly higher in L1 essays than in ELL essays.

An analysis of how each of the significant features was used in the context of ELL and L1 essays revealed the following: Both student groups were still acquiring the appropriate use of modal *would;* the majority of students in both groups did not utilize *it-*clefts; the lower type/token ratio in ELL essays meant that these students used a more limited vocabulary than did L1 students; and ELL students were still acquiring the accurate and appropriate uses of perfect aspect, passive voice, and reduced adjective clauses, whereas L1 students used these features grammatically and for the standard uses.

To apply these findings to the ELL writing classroom, instructors should help students raise their awareness of these six features in their own academic writing by leading students in identifying grammatical and ungrammatical uses of these features and providing practice in differentiating between uses which are standard to the register of academic writing and uses which are appropriate only in conversation. Two sample activities are included to illustrate how to implement these recommendations.

Dedication

To my biggest supporters, Jack and Annette Russell, for shaping the person I've become with their unconditional love.

In loving memory of Kevin Allen Gibler (1984-2014)
A true friend and gentle spirit whose presence in my life has made all the difference.

Acknowledgments

In acknowledging those who have influenced and aided me throughout the process of writing a master's thesis, there are simply too many to name to whom I owe my gratitude. Above all, I thank my parents, John and Annette Russell, for their understanding and patience as I stressed through each deadline, and for their life-long encouragement of me to pursue happiness, whatever that may look like. In that vein, I also thank Sara Marruffo, who has sacrificed more than anyone should have for me to return to school, and whose years of good advice, thought-provoking conversation, and stalwart support have helped me get to where I am today.

I could never have attempted anything like writing a thesis without the guidance of my gracious advisor, Susan Conrad; her encouragement, insight and enthusiasm for this project kept me in perpetual motion toward my goal. I thank her for investing her time in me and for believing in my potential. Thanks to my committee, Susan, Lynn Santelmann and Linnea Spitzer, for generously sharing their time and collective wisdom with me. It is also important that I thank Monika Mulder at the IELP Learning Center, who has been a valued mentor to me, and all of my professors at PSU, who have expanded my mind in ways I didn't know possible.

Finally, my personal gratitude goes to my friends/classmates/colleagues in the MA TESOL program, who are always willing to lend an ear/ commiserate, share their advice and unique perspectives with me, and who have made this "two-year" pursuit of education and enlightenment the most rewarding and enriching experience of my life.

Table of Contents

List of Tables

List of Figures

Chapter 1— Introduction

The purpose of this study is to better understand the writing needs of advanced level English language learners (ELLs) who are preparing to enter degree-granting programs at U.S. universities. Specifically, the research compares the frequencies of linguistic features in the writing of students in main-stream university courses who speak English as their first language (L1) with those of ELL students in an advanced level ELL writing course. The results of the analysis inform decisions about which features to target for pedagogical intervention in ELL writing instruction. In particular, the results further the goal of English for Academic Purposes (EAP) to understand the difficulties specific to ELL students learning academic writing, and inform EAP instructors at Portland State University of any gaps that exist between the writing of ELL students and that of their future L1 classmates.

Writing is a crucial skill; in higher education, so much of academic success is dependent upon the ability to clearly communicate one's ideas through writing. Beyond the university, strong writing skills are sought after in the job market regardless of the profession. However, English language learners face many challenges when learning to write in a second language beyond mastering grammatical structures and amassing a broad vocabulary. ELLs must learn to distinguish between academic and conversational English with regards to the conventional words, phrases, and sentence structures (register knowledge; Biber and Conrad, 2009, Chapelle, 1998, Gardner, 2012). They must also learn the requirements for writing in a variety of genres, including the appropriate information to include and in which order to include it (genre knowledge; Swales, 1990;

Biber and Conrad, 2009; Martin & Rose, 2008). Finally, they must understand the

differences between writing for the academy (skills taught in general EAP courses), and

writing for their specific careers (Gilquin, 2007).

As an aspiring teacher of writing for EAP, I am interested in learning more about

how to prepare students for the type of writing assignments they will be expected to do

for content courses in their degree programs. Through my experience as a graduate

student in the field of Teaching English to Speakers of Other Languages (TESOL), I have

learned about the challenges that ELL students face in English-medium content courses.

For instance, I learned of a Portland State professor of a 100 level course who approached

an EAP instructor to complain that the ELL students were such poor writers compared to

the L1 students that he had difficulty reading their assignments. His claim made me

wonder whether L1 and ELL writing is in fact different, and if so, in what ways? To

begin to answer this, I decided to conduct a quantitative analysis comparing linguistic

features found by prior studies to occur in differing frequencies between these two

student groups.

I chose a corpus-based approach for my research because this approach allows for

the analysis of more linguistic features by more writers than other approaches (Biber,

Conrad, and Reppen, 1998). As I describe in Chapter 2, corpus linguistics provides

empirical evidence for how language is actually used in context, instead of how teachers

or linguists *think* it is used (Granger, 1998, Biber & Reppen, 1998, Breyer, 2011). In

comparison to the findings of earlier approaches, corpus-based studies have provided a

more accurate, descriptive knowledge of English, which has led to the creation of new

materials for enhancing ELL instruction (Reppen, 2012; Chang & Kuo, 2011). Previous

analyses of learner corpora have proven useful for determining the ways that ELL writing

differs from the standard academic register (e.g., Gilquin, Granger, & Paquot, 2007) and

from the writing of L1 students (e.g., Hinkel, 2002). Past corpus-based research of learner

writing has shown that although there is overlap in the sort of inconsistencies that appear

in ELL and L1 student writing compared to texts written in the standard academic

register (Beck, 2007), writing of ELL students contains distinct differences with regards

to the frequency of certain linguistic features in comparison to novice L1 writing

(Gilquin, et al., 2007; Hinkel 2002). One particular way that ELL student writing differs

from L1 student writing is that ELL writing has been found to share more features of the

spoken register of English (Hinkel, 2002; 2003a). Due to differences like this, Gilquin, et

al. (2007) called for more empirical research that utilizes learner corpora in order to

better understand the needs of ELL students and to develop teaching materials that target

specific problems. The present study seeks to better understand the characteristics of ELL

writing in order to inform EAP instructors of the challenges that most need to be

addressed in writing courses and to inform the development of pedagogical interventions

for helping students improve their writing.

My study follows from research done by Hinkel (2002), which compared the

frequencies of 68 linguistic features in a corpus of 1457 essays written by L1 and ELL

students in U.S. universities. I employed Hinkel's method of data analysis but targeted a

more specific set of data: 30 essays each from ELL and L1 students at Portland State

University were analyzed for 12 features found in Hinkel's study to be used in

significantly different frequencies between L1 and ELL writing samples. An additional

feature for counting the diversity of words in a text, type/ token ratio, was analyzed to

compare the breadth of vocabulary used by each group of students. I counted the tokens

of each feature in each essay and performed a statistical analysis to find significant

differences in the frequencies of each feature between the two groups of essays. After

discovering which features were significantly different in frequency, I analyzed the

function(s) of each of these features by comparing examples in the ELL and L1 essays.

Overview of the Thesis

Before describing the present study in detail, I provide the background necessary

for understanding the research questions. Chapter 2 introduces literature about the

challenges that ELL students face in learning to write academic English, the importance

of register as a variable in text analysis, and how corpus linguistics is used to study

second language writing. It then continues with a review of prior research analyzing the

differences in use of linguistics features between L1 writing and ELL writing. Chapter 3

describes the context for the present study and the methods I employed to carry out the

analysis of the selected linguistic features. Chapter 4 presents the results of the frequency

analysis and discusses the findings through an analysis of each feature's use in the context

of the essays. Drawing on the results of this study, Chapter 5 provides suggestions for the

explicit teaching of the significant linguistic features in EAP writing courses. The chapter

concludes with a discussion of the study's limitations and suggestions for future research.

Chapter 2— Literature Review

Teachers and researchers have noted that English language learners face many challenges with the kinds of writing required in college coursework, yet writing is a necessary skill for achieving success in college and in a career. A fundamental part of learning to write for academic purposes is becoming familiar with the genres and registers of academic English, knowledge which is developed over the course of a four-year degree program. Understanding the extent to which ELL students employ the linguistic features of academic registers in their writing is a first step in closing any gaps in ability that may exist between ELL students and their English L1 classmates in university courses. The literature review begins with a discussion of the need to address ELL academic writing and provides information about ELL students in U.S. universities. The sections that follow provide an overview of register variation, corpus linguistics, and past research analyzing the differences in the use of words and structures between ELL and L1 student writing. After that, I discuss the findings of Hinkel (2002), which used a corpus of student writing to compare the frequencies of linguistic features in L1 and ELL writing samples. Finally, I present the research questions addressed in the present study.

The Need to Address ELL Writing

Historically, large numbers of international students have chosen to study in the U.S. due to a lack of educational options in their home countries, the status that a U.S. education carries abroad, and the high quality of education that opens doors for career opportunities in their home countries, the U.S., or a third country (Terzian & Osborne, 2006). Since the 1950's there has been a steady and, at times, exponential increase in the

number of students in U.S. institutions of higher learning who speak English as a second

or additional language. In the 2012-2013 school year, a record number of 819,644

international students studied in U.S. colleges and universities, making up 3.9% of all

students enrolled in U.S. institutions of higher education. This represents a 40% increase

since 2002-2003, only a decade earlier (Institute of International Education [IIE], 2013).

The number of students enrolled in ESL programs in U.S. universities also continues to

grow, with 29,603 students in 2010-2011; this was a 24% increase from the year before

(Smith-Barrow, 2013). At Portland State University, where the present study was

conducted, 10% of 2010 graduates had participated in English as a second language

classes. This represents the university with the largest number of English language

learners in the country for that year (Hopkins, 2011).

The need to address ELL writing stems from the expectations of U.S. institutions

that all students demonstrate a high level of written communication skills by the time

they graduate, whether English is their first or second language. To ensure that ELL

students achieve this level of proficiency, they are placed either in English language

courses to learn the conventions and expectations of university-level discourse, or having

already attained a certain level of language ability, in English composition or general

studies courses enrolling both L1 and ELL students. The approaches of English

department faculty and English as a second language faculty usually differ in how they

offer writing instruction. Often the focus of the former is writing for the humanities and

has a more ideological approach, whereas the latter more commonly focuses on English

for Academic Purposes (EAP) and takes a more pragmatic approach (Santos, 1992;

Benesch, 1993). Because mainstream composition courses are usually offered within English departments, the approach to writing instruction often emphasizes personal voice in writing, as well as the social and political implications of the written word (Santos, 1992). In contrast, second language writing is considered a subfield of applied linguistics, is more product-oriented and values the field's "prevailing standards of inquiry and research" (Santos, 1992, p. 164). The courses tend to focus on the standard expectations for different types of writing (e.g. expository writing or research papers) and emphasize the requirements of an assignment over its wider social and political implications. In any case, ELL students will receive some form of general writing instruction before entering into courses for their major, where academic writing skills will be essential.

Once enrolled in a degree program, ELL students are generally held to the L1 standard of written discourse whether or not they plan to publish research or pursue a writing-intensive career. This expectation has been criticized by those in the field of critical applied linguistics for heralding "native-like" language abilities as a one-size-fits-all standard for ELLs to emulate, while not taking into account the identity, life circumstances, or personal goals of the individual (Ramanathan & Atkinson, 1999; Benesch, 1993). On the other hand, the dominant argument in favor of English for Academic Purposes (EAP) is that while "native-like" abilities are not the end goal, ELL students still must achieve a level of proficiency in writing that will allow them to participate in academic discourse and assert themselves as competent professionals in their field (Santos, 2001; for a discussion of this issue, see Benesch, 2001, Chapter 3 and Casanave, 2004, Chapter 6). A further argument of EAP contends that students of English

as a second or foreign language expect their instructors to teach them the standard

language and conventional organization for college-level writing, with the understanding

that this base knowledge of academic English is vital for their future studies and future

professions. The current study assumes the point of view of EAP, in which the writing

instructor is responsible for providing a foundation upon which ELL students may build

the language and communication skills needed for the writing tasks that they will do in

mainstream university courses.

As mentioned previously, an important aspect of becoming proficient in academic

English is learning to write appropriate academic language and becoming familiar with

the genres of academic and professional life. Both L1 and ELL students build this

knowledge of academic discourse through reading the work of authorities in the field and

through the practice of writing the sub-varieties of academic English, such as the

persuasive essay and the lab report. While teaching students about the sections required

in a lab report is rather straightforward, teaching which words are most effective for

expressing the ideas within the lab report is much more challenging since even teachers

of English language are often unaware of the choices they make regarding the words and

grammatical structures in their own writing. My experience has been that EAP teachers

may unwittingly expect that if ELL students work hard enough in their various EAP

courses, their writing will eventually begin to take on the characteristics of the academic

language, but in fact, students could greatly benefit from receiving explicit instruction in

this area prior to entering content courses.

In fact, research implementing the genre-based approach to instruct ELL students in academic writing has found that after explicit instruction of genre and register (organization of ideas and linguistic features, which will be explained in the next paragraph) students were better able to produce essays that followed the conventions of academic writing (Kongpetch, 2006; Cheng, 2008; Chen & Su, 2012; Yasuda, 2011; Yayli, 2011). For example, Kongpetch (2006) used this approach to introduce Thai university students to persuasive essay writing. The instruction consisted of four steps: building students' content knowledge, modelling the organization and linguistic features used in the persuasive essay and the rhetorical moves within and between each paragraph, co-constructing persuasive essays in small groups, and finally, independently writing persuasive essays, with students putting into practice what they had learned in the first three steps. The final product, students' own persuasive essays, provided evidence that students were able to follow explicit instruction to successfully produce writing that contained the appropriate organization of ideas and linguistic features that reflect the academic written register.

Up to this point in the literature review, I have used the terms *genre* and *register* without examining the differences between the two. Since previous studies have sometimes used *genre* and *register* interchangeably, and some researchers have changed their operational meanings over time (Biber & Conrad, 2009), it is worth clarifying the difference between them before explaining the importance of register to my study.

Distinguishing between Genre and Register

The terms *genre* and *register* both apply to analyzing characteristics of varieties of texts, but when the terms are clearly distinguished and defined, they are typically differentiated in the following way (e.g. see Martin & Rose, 2008; Gardner, 2012; Chapelle, 1998; Biber & Conrad, 2009; Hyland, 2000): Genre characteristics refer to the sequencing of information that is characteristic of the type of text, whereas *register* refers to the distribution of linguistic features that occur throughout the text type. An example of genre is the IMRD structure of a lab report (Introduction, Methods, Results, Discussion); the term also refers to once-occurring or rare linguistic features that are characteristic to a type of text, such as starting a letter with "Dear [name]." On the other hand, a study of register features might find that the relative frequency of passive voice is higher in research articles than in personal letters, while the personal letters have a higher frequency of personal pronouns.

While the focus of genre analysis and register analysis differs, both aim to provide a better understanding of the function of language in a particular context. For example, a study of genre might analyze research articles for execution of the IMRD structure and would acknowledge the function of this structure: to make very detailed and complex ideas understandable to a large audience of readers through delineating the steps of a research project in chronological order, from providing necessary background information about the topic (introduction), the methods used to perform the experiment, the results of the experiment, and finally, a discussion of why the results matter. The analysis would also highlight how any deviations from the structure affected the

readability of the text. Similarly, the register analysis would account for the findings of

more non-passives and pronouns in letters by explaining that letter writers are likely to

refer to themselves and their audience directly. While it is helpful from a pedagogical

standpoint to understand to what extent ELL students employ genre conventions, for the

present study I chose to focus attention on register knowledge, specifically the

distribution of linguistic features in ELL writing, because I believe it will provide insight

into an especially challenging aspect of learning to write in a second language at the

university level. In the next section I will further explain how the concept of register

relates to the present study.

Register Variation

The present study incorporates the concept of register in that it seeks to discover

whether ELL students, when compared to L1 students, use certain linguistic features

more or less frequently in their academic papers. University students in the U.S.

regardless of first language learn academic registers throughout their education by

reading published works and by writing in-school genres such as essays and research

reports. However, the difficulty of learning these registers can be illustrated by the fact

that many people who speak English as a first language never become proficient readers

or writers of academic papers. ELL students may struggle even more with learning these

registers because they have not been gradually exposed to academic English throughout

their secondary school education, as L1 students have. In addition, ELL students in U.S.

universities are regularly exposed to the conversational register of English and may

incorporate their burgeoning knowledge of spoken English into their academic writing.

Unless they have been directed by a teacher to do so, they likely have not thought about the situational context, purpose, and audience of academic writing and how it differs from that of face-to-face conversation. Even if students have a sense that everyday conversation and academic English differ, they may not have been taught which linguistic features to use and which to avoid when writing a paper for class. The importance of teaching ELL students about registers, which includes situational contexts, linguistic features, and their associated functions, is the pedagogical motivation for conducting the present study.

As mentioned before, registers are distinguished by the situations for which they are intended, as well as by the relative distribution of linguistic features they employ; a particular register will typically contain a higher frequency of certain features, while another register will have a lower frequency of these same features. The linguistic features analysis at the center of the present study relates to the study of register variation, in that it will examine and compare the distribution of features in the academic writing of the two learner groups, L1 and ELL students. This type of analysis requires the use of corpus linguistics, the branch of applied linguistics that has made it possible to analyze large samples of spoken and written texts with the goal of describing how language is actually used. The following section discusses corpus linguistics, and highlights past research that has utilized corpora to analyze the distributions of linguistic features in ELL and English L1 student writing.

Corpus Linguistics

Corpus linguistics involves the use of a corpus, which is a large, principled collection of written and/or transcribed spoken texts, for large-scale analyses of naturally occurring discourse. Some corpora contain millions of words from multiple registers of English (Davies, 2008-), while others contain fewer words and focus on a smaller number of registers (Granger, 2003). Due to its capacity to efficiently analyze large numbers of texts, corpus linguistics has advanced the fields of linguistics and English language teaching in important ways. First, it has allowed for an empirical analysis of language, meaning that what is deemed appropriate language for a given context is informed by examples from authentic language use instead of by human perceptions (Biber & Reppen, 1998). This has resulted in an increasingly more detailed and accurate understanding of how English language is actually used (Granger, 1998). What's more, empirical analysis of corpora "has enabled researchers to discover patterns of language usage that had previously remained hidden from view" (Breyer, 2011, p. 1). These patterns are analyzed not only quantitatively, but also qualitatively, as corpus linguistics also entails the analysis of the specific functions performed by language. The deeper linguistic knowledge generated by corpus studies has influenced the creation of new dictionaries (Biber et al., 1999; Gillard & Gadsby, 1998), textbooks (Reppen, 2012), and teaching materials (Chang & Kuo, 2011) that aim to provide students with an accurate and descriptive knowledge of the English language.

Evidence that corpus methodologies have advanced the field of linguistics can been seen in register analysis, which relies on corpus methodologies to better understand

the patterns of English used in a given register. In a study which utilized corpora to study

the variation between registers, Biber (1988) analyzed the co-occurrence of 67 linguistic

features to create continuums of variation, called dimensions, which predict the

differences between spoken and written registers. For example, on a continuum between

spoken and written English, a higher type/token ratio (the number of unique words in a

text divided by number of total words) corresponds with a lower frequency of private

verbs, emphatics, amplifier adverbs, *be* as a main verb, and coordinating conjunctions,

and vice versa. Biber explains that the latter group of features is indicative of interactive

communication with a non-informational focus, while type/token ratio, co-occurring with

several other features, is indicative of a text with a high informational focus, a great

density of information, and the communication of information in a concise and precise

way (pp. 104-5). Thus, academic writing typically demonstrates a higher type/token ratio

because writers have had time to plan their ideas, choose their words carefully, and revise

their choices to ensure that the words effectively convey the nuances of their ideas. In

contrast, conversation has a lower type-token ratio since the constraints of on-line

production mean that the speaker tends to use more commonly used words (Biber, 1988).

These findings help explain how registers are dependent on the relative frequency of a

co-occurring set of linguistic features. This study also illustrates how corpus-based

studies can help explain patterns in language that are not obvious to the naked eye.

    Corpus-based studies have compared the distributions of linguistic features in

textbooks and published research articles (Conrad, 1996), student-produced versus

professionally-produced texts (Shaw, 2009), and the writing of L1 versus the writing of

ELL students (Hinkel, 2002). While corpus linguistics takes a variety of approaches to

analyzing language, the approach most pertinent to the present study is as follows:

Researchers compile and tag a corpus of text samples and employ a concordancer to find

all the examples (tokens) of a specific word or linguistic feature in the corpus. They then

calculate the frequency rate of the feature in each text and perform inferential statistics to

compare the frequency rates between the groups of texts (Gilquin, Granger, & Paquot,

2007). Finally, the comparison of frequencies is combined with an analysis of how each

group employs the feature in context, revealing the similarities and differences between

the groups regarding how they understand the functions of the linguistic features. The

applications of the findings of a corpus-based study depend on the text type being

analyzed; when the analyzed texts are produced by university students (whether ELL or

English L1), corpus-based research provides implications for which features to teach or

reinforce at the university level, where students are expected to produce writing that

increasingly meets professional standards. In the next section, I first introduce the

different approaches used to study second language writing, and then focus in on corpus

studies that compare the frequency of linguistic features in academic writing produced by

ELL and L1 students.

Studies of Second Language Writing

There are many approaches to studying effectiveness in second language writing.

They include: identifying the metacognitive processes that ELL students employ while

writing and the strategies for helping students build metacognitive awareness (e.g. Ong,

2014; Yayli, 2011; Negretti & Kuteeva, 2011); measuring cohesion and syntactic

complexity in ELL writing (e.g. Ortega, 2003; Crossley & McNamara, 2012; Kormos, 2011); analyzing theme/ rheme and information structure in second language compositions (e.g. Jalilifar, 2010; Wei, 2013; Rafiei & Modirkhamene, 2012); testing the effects of teacher feedback or peer feedback on subsequent drafts of written work (e.g. Diab, 2010; Evans, Hartshorn, & Strong-Krause, 2011; Nurmukhamedov & Kim, 2010; Bitchener & Knoch, 2010); and exploring use of collocations and metaphor in ELL writing (e.g. Laufer & Waldman, 2011; Kathpalia & Carmel, 2011).

The approach at the center of this study is analyzing frequency and function of linguistic features in student writing to find where they match the standard uses for academic writing and where they deviate. Most studies analyzing linguistic features in student writing utilize corpus methodologies; the rest of this section focuses on the results of these studies.

Employing the corpus-based approach to counting and comparing use of linguistic features (described in the last section), prior studies have found that L1 and ELL writing differs in many respects, including use of collocations and lexical bundles (Siyanova & Schmitt, 2008; Chen, 2010), cohesive devices (Crossly & McNamara, 2009; Hinkel, 2001), adverbial connectors (ie. *for example, however*; Altenberg & Tapper, 1998), and hedging modals (Hyland & Milton, 1997). As the present study analyzes a restricted set of linguistic features, the discussion of past research will focus on what is known about how these particular features are used in ELL writing: *be* as a main verb, predicative adjectives, private verbs and public verbs, *it*-clefts, type/token ratio, passive voice,

perfect aspect, modal *would,* amplifying and emphatic adverbials, and modal verbs of necessity (the features are in bold below; refer to table 2.1 for examples of each feature).

Features of spoken English. In Hinkel's corpus-based analyses comparing the writing of L1 and ELL university students, a consistent finding is that the features employed by ELL writers are more akin to the spoken register than to the academic register, and that the use of these features lends a sense of simplicity and lack of quality to ELL writing. Features of the spoken register include ***be* as a main verb**, **predicative adjectives**, **private verbs** and **public verbs**, all of which were found in statistically higher median frequency rates in the writing of ELL students compared to English L1 essays (Hinkel, 2003a). In contrast, an alternative sentence structure common to academic prose, the ***it*-cleft**, was found in the same study to be statistically more frequent in L1 writing than in ELL writing. Hinkel concluded that despite having received more college-level writing training than their L1 peers, the ELL writers employed less variety of syntactic structures and had a more limited lexicon. She suggested that ELL students receive more focused instruction in how to employ the features of the academic register.

Lexical diversity. Connected to the idea of limited lexicon is the feature **type/ token ratio,** which is a measure of lexical diversity and lexical specificity. Past research has shown that compared to L1 writing, ELL writing has a lower type/token ratio, which corresponds to a smaller range of vocabulary and the use of less specific words (Kormos, 2011; Crossley & McNamara, 2009). Therefore, the inclusion of

Table 2.1

Examples of Linguistic Features and Findings of Prior Research Regarding Frequency of Features in Student Writing.

Feature	Examples	Findings of Prior Research
Public verbs	agree, mention, present	Higher frequency in ELL writing; associated with conversation (Hinkel, 2003a)
Private verbs	accept, consider, find	Higher frequency in ELL writing; associated with conversation (Hinkel, 2003a)
Amplifier adverbs	always, even (+adj), very, well	Higher frequency in ELL writing (Hinkel, 2003b); heightened feeling associated with conversation (Biber, 1988); ELLs overstate the importance of an assertion (Lorentz, 1998)
Emphatics (adverbs and adjectives)	complete, completely, indeed, such a (+noun)	Higher frequency in ELL writing (Hinkel, 2003b); heightened feeling associated with conversation (Biber, 1988); ELLs overstate the importance of an assertion (Lorentz, 1998)
Be as a main verb	*the wife **was** pregnant*	Higher frequency in ELL writing; associated with conversation (Hinkel, 2003a)
Predicative adjectives	*the wife was **pregnant***	Higher frequency in ELL writing; associated with conversation (Hinkel, 2003a)
Necessity modals	must, have to, should, need to, ought	Higher frequency in writing of East Asian ELL students than in English L1 student writing (Hinkel, 2009); same found in Swedish ELL students Aijmer (2002); attributed to differences in cultural values and a more direct style of persuasion.
Modal *would*	*they **would** stand up*	ELL students underused it; need it for depersonalized or objective writing (Hinkel, 2004)
Perfect aspect	*frogs **have been found***	ELL students underused it; need it for depersonalized or objective writing (Hinkel, 2004)
Passive voice	*children **are affected***	ELL students underused it; need it for depersonalized or objective writing (Hinkel, 2004)
Reduced adjective clauses	*The authors **listed in the citation** were students*	Higher frequency in L1 writing (Hinkel, 2002)
it-clefts	***It is the senior women who** keeps track of them*	Higher frequency in L1 writing; a more advanced syntactic structure (Hinkel, 2003a)
Type/token ratio	# of unique words in first 400 words of text	ELL writing has lower measure of lexical diversity and specificity than L1 writing due to less developed word knowledge (Kormos, 2011; Crossley &McNamara, 2008)

type/token ratio in this study provides an additional dimension for better understanding

the differences between the texts produced by the two student groups.

Tense and aspect. Another study by Hinkel (2004) compared the use of English

tense and aspect across seven L1 groups. It was found that ELLs from East and Southeast

Asian countries overused the past tense compared to English L1 students, often

conveying their arguments through personal narrative. Arabic L1 students, on the other

hand, used significantly fewer past tense verbs than English L1 students, using the

present tense to make arguments based on sweeping generalizations and metaphors about

common knowledge not directly related to the essay prompt. All groups of ELL writing

underused or virtually ignored complex verb structures like **passive voice**, **perfect aspect**

and use of modal *would*. These findings highlight the difficulties of ELL students in

producing academic writing that sounds appropriately depersonalized and objective, even

after they have received academic writing instruction. On the other hand, Hinkel (2004)

found that even the most basic L1 writing demonstrated a wider grammatical range and

employed the present tense significantly more frequently than all except the Arabic ELL

group, using present tense for generalizations, observations, and descriptions in

accordance with the conventions of academic prose.

Adverbial markers. In a study of adverbial usage in L1 and ELL writing, Hinkel

(2003b) found that ELL students of all L1 backgrounds in the study used **amplifying** and

emphatic adverbials at higher median frequency rates than freshman L1 university

students. These adverbials have an interpersonal function, marking "heightened feeling,"

and thus are more frequent in spoken than written registers (Biber, 1988, p. 106). One

explanation for the overuse may be that ELL students in U.S. universities receive more

English input through conversation than through academic prose. The relative overuse of

emphatics and other intensifier adverbs in ELL writing was similarly found in a corpus-

based study comparing the writing of German and British students at both the high school

and university levels (Lorenz, 1998). German high school students used the highest

frequency of intensifier adverbs, followed by German university students, then British

high school students, and was used least by British university students. The author argues

that the overuse in ELL writing was likely due to a lack of experience in academic

writing which led ELL students to overstate the importance of an assertion. L1 and ELL

use of intensifiers also differed in where they occurred in the sentence: ELLs tended to

use adverbs as intensifiers in the theme of a sentence, which led to a heavy subject noun

complex, while L1 intensified the adjectives in the more reader-friendly rheme position,

"where one would expect to find the elements that are new, relevant, and noteworthy

enough to be intensified" (Lorenz, p. 62).

        Modal verbs of necessity. Research about the effect of topic on the frequency of

modals in L1 and ELL writing found that high-level ELL students from China, Japan, and

Korea used significantly higher frequency rates of **necessity and obligation modals**

(*must, should, ought to, have to, and need (to)*) than did L1 students in four of six essay

prompts given (Hinkel, 2009). All four of these topics prompted students to discuss

socio-cultural values about roles of family members, while the other two topics relied less

on cultural background. Hinkel attributed the higher rate of necessity modals in family-

related topics to the rigid hierarchical family structure of East Asian cultures influenced

by Confucianism. Another study of modal verb usage found similar results: Compared to

English L1 students in British universities, Swedish L1 university students overused

necessity modals in their academic English writing, "adopting a direct and emphatic

style of persuasion" (Aijmer, 2002, p. 65). Additionally, the ELL and the English L1

students used *must* in different ways. While both groups of writers similarly used *must*

for conveying logical necessity, *must* as personal obligation was found in ELL writing

four times more frequently than in L1 writing. Usage also varied by topic, with a high

number of *should* and *must* found in ELL essays about immigration and environmental

policies; in these cases, the modals were used to convince the reader of the merits of a

certain action, reflecting the "cultural norms of behavior or a moral code" of the writer

(p. 65).

This section of the literature review presented an overview of past research that

found differences between L1 and ELL writing with regards to the use of the linguistic

features related to the present study. In corpus-based studies conducted by Hinkel and

others, ELL writing has been found to contain higher frequencies of features related to

conversation. These features include *be* as a main verb, predicative adjectives, private

verbs and public verbs, and amplifying and emphatic adverbials. Higher frequency of

necessity modals in ELL writing was attributed to cultural transfer. Meanwhile, features

found to be underused in ELL writing are characteristic of the academic register: *it*-clefts,

passive voice, perfect aspect and modal *would*. A high type/token ratio is also

characteristic of academic text with an informational focus. In the final section of the

literature review I introduce Hinkel (2002), a large-scale analysis of linguistic features in

student writing which provides a basis for the current study.

Hinkel's (2002) Study of L1 and ELL Writing

Hinkel (2002) compared frequencies of 68 linguistic features in a corpus of 1,457

essays written by English L1 and ELL students from six L1 backgrounds studying in U.S.

universities. The L1 writers were all first-year students enrolled at public and private

universities in the U.S. They chose one of six prompts designed to elicit an argumentative

or expository essay; the mean number of words per essay was 294. The essays were

written as a diagnostic test for placement into first-year composition courses. Thus, none

of the L1 writers had received college level writing instruction at the time they composed

the essays, and their experience with academic writing was assumed to be limited to what

they had learned in high school English classes.

ELL writers were given the same essay prompts as the L1 writers for ease of

comparison; the mean number of words per essay was 299. They all had a high level of

English language proficiency as demonstrated by their TOEFL scores. All were students

at four-year universities in the U.S. and were enrolled in courses in their major at the time

of the study. They had previously completed English language training in intensive

English programs either in the U.S. or abroad, or in community college or in high school

ESL programs in the U.S. The majority of them had transferred to their current

universities as juniors after earning their associate's degrees from two-year community

colleges. The rest were either first-year students or graduate students. As a group, the

ELL writers had received a mean of 9.8 years of EFL or ESL instruction. Therefore, they

had had more college-level academic writing experience than the writers in the L1 group.

The results of the analysis found 26 features that were used in significantly

different frequencies between the L1 and ELL groups. Sixteen features associated with

simplistic writing were found in the ELL texts at a median frequency rate statistically

higher than that of the English L1 writing samples (Table 2.2); they included *be*-copula

as a main verb, predicative adjectives, and public and private verbs. Ten features were

more frequent in L1 writing samples, including passive voice, perfect aspect, and reduced

adjective clauses. The 42 features that were not found in significantly different

frequencies included personal and indirect pronouns, existential *there*, verb tenses (past,

present and future), and modals of possibility and ability.

The results revealed that the writing of English L2 university students trained in

academic writing was significantly less syntactically and lexically complex compared to

the writing of native English speakers who had not received instruction in college-level

writing. Hinkel concluded that by educating students of the pitfalls of simplistic writing,

teachers could help students notice these features in their writing and correct them on

their own using the strategies they had been taught. She provided practical examples for

better preparing students to achieve the level of writing ability expected of them in their

university courses. One such example was to present ELL writers with alternatives to the

most commonly used public and private verbs.

While it is helpful to know which features ELL students used in lower or

Table 2.2

Features of Text Found in Significantly Different Frequencies between Samples of L1 and ELL Writing (Hinkel, 2002)

Classes of features	Features with significantly higher frequency rates in ELL texts	Features with significantly higher frequency rates in L1 texts
Nouns	interpretive nouns vague nouns assertive pronouns	*it*-cleft
Verbs	public verbs private verbs expecting/ tentative verbs modal verbs of necessity *be* as a main verb	perfect aspect progressive aspect predictive modal *would* passive voice present participles past participles
Adjectives/ Adverbs	predicative adjectives amplifiers other adverbs (manner, conjunct, and adjective/verb modifiers) adverb clauses of cause	
Subordinate clauses		reduced adjective clauses reduced adverb clauses
Text-rhetorical features	phrase-level conjunctions sentence-level conjunctions (transitions) exemplification markers (for example) emphatics	fixed strings (idiomatic phrases and collocations)

higher frequencies compared to L1 students, a limitation of Hinkel's study is that the features with significantly different frequencies were not examined more in depth to discover and compare the situations and meanings for which the two groups of writers employed these features in the context of their essays. The present study addresses the need to better understand the functions of the features employed in student writing in cases where L1 and ELL students' frequency of use is significantly different.

Summary and Research Questions

The expectations of the academy and of students themselves make ELL writing an area of instruction and research that deserves attention if we as educators aim to maximize our opportunity to prepare students for academic success. The literature review introduced the importance of register knowledge in learning to write academic English and explained how corpus linguistics has been applied to study the frequency and use of linguistic features. Corpus-based studies have found many differences between L1 and ELL academic writing, but a major finding is that ELLs use a greater frequency of features associated with conversation, while using a lower frequency of those associated with academic prose. The present study is based on Hinkel (2002), but builds on those findings to better understand ELL writing in the specific context of Portland State University's Intensive English Language Program, especially how it differs from L1 writing with regards to the frequency and function of linguistic features. Specifically, this study will address the following research questions:

1. Does the writing produced by English L1 undergraduate students and ELL students of English for Academic Purposes contain significantly different frequencies of the linguistic features found to be different in Hinkel (2002)?

2. For features with significant differences, in what ways do ELL and L1 students use these features differently in the context of their writing, and what do these differences reveal about students' academic writing skills?

The importance of the frequency differences will be interpreted with additional analysis of the features' use in order to provide suggestions for teaching academic writing.

Chapter 3—Methods

This chapter describes the methodology of the study. The first section presents the context of the study; it describes Portland State University's undergraduate programs and the Intensive English Language Program, the two settings in which the analyzed essays were produced. The second section details the data collection process, and the third section describes the data analysis procedures.

Context of Study

This study, conducted at Portland State University (PSU) in Portland, Oregon, compared the writing produced by English L1 undergraduate students enrolled in a variety of courses in University Studies and 9 other departments, and the writing produced by English language learner students enrolled in Level 5 writing courses in the Intensive English Language Program (IELP) at PSU. Below I describe the various courses and programs through which students at Portland State University learn academic writing.

Writing at PSU. Portland State University is a public university in an urban setting. It is the largest university in Oregon, with nearly 30,000 students and over 200 degree programs from bachelor's to doctoral degrees ("Portland State University Facts," n.d.). Despite the university's size, the average class size in freshman courses is 24 students ("Profile," n.d.). While students come from all 50 states and 80 countries, 79% of students are Oregon residents. PSU is the most racially diverse university in the state,

with non-white students representing 38% of the student body. It is also economically

diverse, with 59% of students receiving financial aid ("Profile").

To ensure that students learn effective communication skills, the university has a

writing requirement for all students pursuing a bachelor's degree ("NEW: University

Writing Requirement," n.d.). For students entering as freshmen, the requirement is

fulfilled through mandatory enrollment in a sequence of University Studies courses (this

program is described below). Transfer students must take a combination of University

Studies courses and composition courses (in the English department), or writing intensive

courses (in other departments; "NEW: University Writing Requirement"). For example,

students transferring into the Maseeh College of Engineering and Computer Science with

between 30-59 credits are required to take Writing 121 College Writing, Communication

100, and six University Studies courses for fulfilment of the college's Arts and Letters

and Social Sciences requirement (Maseeh College of Engineering and Computer Science,

2011). Thus, the primary avenue through which PSU students learn about academic

writing is the University Studies program.

University Studies. University Studies is the general education program required

for undergraduate students at Portland State University (except for Honor's College and

some transfer students). According to the university website, the purpose of the program

is to broaden students' knowledge of the diversity of human experience, ethics and social

responsibility while improving upon the academic skills required at the post-secondary

level ("University Studies Goals," n.d.). University Studies entails a year-long course in

the first year, three term-long courses each during the sophomore and junior years, and

the completion of a senior capstone course in which students contribute to the community

outside the university: the Portland metropolitan area ("UNST Introduction," n.d.).

Examples of courses in the University Studies program are "Race and Social Justice" and

"Sustainability." Each course aims to develop the inquiry, critical thinking, and writing

skills needed for success in higher level college courses. Students produce writing that

represents several genres and sub-genres, including literary analysis, secondary research

papers, reactions to a reading, and reports of first-hand experiences.

IELP. The Intensive English Language Program (IELP) is a 6-level academic

English preparatory program that provides instruction in English language and study

skills to prepare international students for success in undergraduate and graduate degree

programs at U.S. institutions. In the fall of 2012, the first term that the ELL essays were

collected, over 500 students were enrolled in the program. Of these, 53% of students

were from Saudi Arabia and 6% were from Kuwait (Arabic L1), 10% were from China

and 3% were from Taiwan (Chinese L1), 10% were from Japan (Japanese L1), 4% were

from South Korea (Korean L1), and the rest were from 34 other countries (IELP, 2013).

The program is operated by Portland State University within the Department of Applied

Linguistics. Before students begin taking courses in the IELP, they take a placement test

which assigns them to a general ability level; if initially placed in the lowest level, it will

take 6 terms, or 1.5 school years, for them to complete the intensive English program,

provided that they achieve passing grades in every course. The lowest level is the Pre-

entry Program (PEP), followed by levels 1-5. At each level students are required to take

Speaking and Listening, Grammar and Writing, Reading, and an elective course
(Intensive English Language Program, n.d.).

The writing courses in the two highest levels of the IELP (levels 4 and 5) are
called Guided Research Writing and Independent Research Writing. The goal of these
courses is to familiarize students with the process of writing a research paper so they will
be prepared to write similar papers in future college courses. While level 4 lays the
groundwork for formatting a research paper using a topic and sources chosen by the
teacher, level 5 expands on these research-paper writing skills by requiring students to
choose their own topic, collect and evaluate sources to ensure they are relevant to the
topic, and create an original research question, culminating in an 8-10 page research
paper (Smith, 2010).

Upon completion of level 5 coursework, IELP students often matriculate into a
degree program at Portland State University and begin taking content courses with
English L1 classmates. However, a full 40% of IELP students begin taking mainstream
courses at PSU while concurrently enrolled in IELP levels 4 and 5 (B. Bolstad, personal
communication, April 19, 2013). These students speak a language other than English as
their first language and have been conditionally accepted into a degree program pending
completion of the university's language proficiency requirement. Conditionally-admitted
freshmen may take one 100- or 200-level PSU course per term while enrolled in IELP
level 4 and 5 courses. Conditionally-admitted IELP students who have transferred to PSU
from another university are allowed to take one 300-level course per term with
permission from an advisor.

Whether preparing to enter PSU courses or already taking them, students in level 5 writing are at a critical juncture in their education between English-language and content courses; it is at this point that ELL students most rely on the skills learned in the IELP. Because ELL students' academic work in PSU courses is compared to the work of their English L1 classmates who have been exposed to academic register of English through years of formal education, it is also at this point when the level of preparation provided by IELP courses becomes most apparent. This study seeks to compare linguistic features in the writing of two student populations that share or will soon share classrooms and professors, and whose writing skills will be in clear relief.

Data Collection

I analyzed thirty essays written by ELL students in the IELP level 5 writing course, and thirty essays by English L1 students written for 100-, 200-, and 300-level courses in the undergraduate departments described above. Essays in the two groups contained between 850 and 4000 words each, with an average length of 1838 words in the L1 corpus and 2220 words in the ELL corpus (see Table 3.1).

English L1 corpus. The samples of L1 writing were selected from the Portland State University Corpus of Student Academic Writing (Viking Corpus), a project within the Department of Applied Linguistics (see Albers, 2007). The Viking Corpus represents papers from many disciplines and from students at every level of the institution; all papers received an A or a B grade and fall into one of the following categories or sub-genres of academic writing, as labeled by the Viking Corpus: art and

Table 3.1

Average Words per Sample and Total Words in the L1 and ELL Corpora

	Corpus of L1 writing	Corpus of ELL writing
Average words per sample	1874	2306
Total words in corpus	56,210	69,178

Table 3.2

Number of Samples from Each Department and Course Level in L1 Corpus

Department	100 level	200 level	300 level	Total
University Studies	9	1	0	10
International Studies	0	1	3	4
Anthropology	0	0	4	4
History	1	2	0	3
Political Science	1	0	2	3
Psychology	0	0	2	2
Women's Studies	0	0	1	1
English	1	0	0	1
General Science	0	1	0	1
Philosophy	0	1	0	1
Total	12	6	12	30

literary analysis, empirical analysis, library research, narrative, personal opinion, reading

reaction, reflection, and report. The corpus compiled for the current study includes 30

papers from the Viking Corpus. To ensure that the writing samples were as comparable to

one another as possible, all papers were of the 'library research' sub-genre, the

characteristic college writing assignment in the which the author cites secondary sources

for information and does not collect original data (Albers, 2007). One-third of the essays

were written for courses in the University Studies program, and the other two-thirds were

written for 100, 200, and 300-level courses in a variety of departments (see Table 3.2).

This collection is meant to represent the variety of courses that IELP students are

currently taking or will take upon completion of the level 5 writing course. The 30

English L1 papers were selected from the Viking Corpus based on the student L1

(English), course level (100-, 200-, and 300-levels), and length of sample (between 800

and 4000 words).

ELL corpus. The papers in the ELL corpus were written by students in the IELP

level 5 Independent Research Writing course. To ensure comparability with the L1

corpus, the papers in this group were also of the library research sub-genre and received

either an A or B grade. This sample was expected to capture the level of writing ability

demonstrated by high-achieving students in the final term of this intensive English

program. The 30 ELL papers were collected by Linnea Spitzer, an IELP writing

instructor, for the purpose of conducting research on the characteristics of student writing.

She received approval from the Human Subject Research Review Committee in the

spring of 2012 (HSRRC # 122234) and began collecting informed consent forms in the

fall of 2012. Six classrooms were represented in the ELL samples: two sections each

from the fall 2012, winter 2013, and spring 2013 level 5 Independent Research Writing

course. Linnea Spitzer was the instructor in three of these sections; the teachers of the

other sections varied, but the specific teacher was not expected to affect the frequency of

features in student writing.

Table 3.3 presents an approximation of the various disciplines represented in the

ELL corpus; because essay topics were chosen by the students, they reflect a variety of

interests, from the humanities to business to computer science. Essay topics sometimes

straddled two disciplines, such as sociology and psychology in the case of an

Table 3.3

Examples of Essay Topics and Disciplines Represented in ELL Corpus

Disciplines of essay topics	No. of essays	Example essay topics
Sociology/Psychology	7	Video game addiction: effects on the individual and society
Business	7	Marketing: the privacy concerns of behavioral targeting
Environmental studies	5	The benefits of composting; Solar energy in Saudi Arabia
International studies	3	The benefits of volunteering abroad; Reverse culture shock
Education	3	The use of corporal punishment in US schools
Public health	2	The public health benefits of Medicaid and recent changes
History	1	The destruction of Kuwait by the Iraqi invasion of 1990
Engineering	1	How architectural design affects quality of life and mood
Computer science	1	Common problems associated with internet browsers
Total	30	

essay about sociological implication of the divorce rate in the U.S. and the psychological effects of divorce on the child. While ELL students wrote mainly about business, the environment, and education, L1 students wrote largely about anthropology, political science, international studies. The similarities between corpora are that both consist mainly of essays written within disciplines in the Liberal Arts, and there is also some overlap of topics (childhood trauma, reverse culture shock).

Data Analysis Procedures

I used concordancer technology and non-parametric statistics to conduct a quantitative analysis of selected linguistic features to determine whether they were used in significantly different frequencies between the 30 L1 and 30 ELL writing samples. The following section outlines the procedures for selecting features to analyze, counting the number of tokens of each feature, obtaining statistical results, and analyzing the use of each feature in context.

Table 3.4

Features Analyzed for Frequency Rate in the Present Study

Type of feature	Semantic classes	Structural classes	Interpersonal classes
Verbs	public, private, 'be' as a main verb	perfect aspect, passive voice	necessity modals, modal *would*
Adverbs/ adjectives		predicative adjectives	amplifiers/ emphatics
Clausal		it-cleft, reduced adjective clause	
Phrasal		phrase-level coordination	
Lexical diversity	type/token ratio		

Selecting features. I selected features from those which were analyzed in Hinkel

(2002). Hinkel's study found 26 features used in significantly different frequencies

between L1 and ELL writing samples (see Table 2.2). For the scope of a master's thesis, I

needed to narrow down the number of features to analyze. I chose a set of features that

represent the variety of different parts of speech and semantic functions that Hinkel

covered without repeating many features in any single category. I also chose the set to

include some features found by Hinkel to be used in greater frequencies by ELL students

and others found to be used in greater frequencies by L1 students. I settled on the 13

features listed in Table 3.4 (for examples of each feature, refer to Table 2.1). Because

verbs are at the heart of the English clause (since a clause is built around the verb) the

analysis included seven different verbs classes: the semantic classes of public verbs,

private verbs, and *be* as a main verb; the interpersonal classes of necessity modals and

modal *would*; and the structural classes of perfect aspect and passive voice. Passive voice

includes all passivized constructions with and without a *by*-phrase, and includes post

nominal modifier –ed participle clauses, which are referred to as *reduced adjective*

clauses in the present study. The selected features also included the following classes of

adjectives and adverbs: predicative adjectives, amplifier adverbs, and emphatics, which

can act as both adverbs and adjectives with an interpersonal function. Finally, clausal and

phrasal features included coordinating conjunctions, reduced adjective clauses (post-

nominal modifiers with past and present participles), and the *it*-cleft, which is a marked

structure characteristic of academic writing. An additional feature not of semantic,

structural, or interpersonal relevance, type/token ratio, is a measure of lexical specificity

and diversity and has been used for comparing the breadth of vocabulary employed in the

spoken and written registers (Biber, 1988), as well as in the writing of English L1 and

ELL university students (Kormos, 2011). Another motivation for my choice of features

was to discover whether the ELL essays in the present study corroborated Hinkel's

finding that ELL writing contained more features typical of conversation than did L1

writing. Therefore, I made sure to include the conversational features found to be more

frequent in ELL writing: *be* as a main verb, predicative adjectives, private verbs, public

verbs, and amplifiers/ emphatics (Hinkel, 2003a; 2003b).

      Coding of features. Each writing sample was grammatically tagged with the

Biber tagger (see Biber, Conrad, & Reppen, 1998; Biber, 2006), which coded each word

by its part of speech and, in some cases assigned a more specific classification for

semantic class or other subcategory distinction. I used MonoConc Pro 2.2 concordance

software (Barlow, 2004) to count each feature. In order to get the most accurate count, I

used the word lists provided in Hinkel (2002) and searched for each of the words

subsumed under the following categories: private verbs, public verbs, amplifiers,

emphatics, modal *would,* necessity modals, *it*-cleft, and phrase-level coordination.

Monoconc Pro produced a list of all of the tokens of a searched word, and I went through

each list to check that each word was in fact an example of that feature. For example,

when I performed a text search for *thought*, the past tense of the private verb *think,* I had

to check each token found by the concordancer to be careful that I only counted the

tokens of *thought* as a verb and not the tokens of the noun *thought*. I counted the

following features by searching for their tags: predicative adjectives, *be* as a main verb,

passive voice, perfect aspect, and reduced adjective clauses. To ensure accuracy, I

searched each document thoroughly for tokens that were mis-tagged, and adjusted counts

accordingly. For example, as I checked that each token of *be* as a main verb was correctly

tagged, I also checked that tokens of predicative adjectives, which co-occur with this

feature, were also all tagged correctly, and included the tokens that had not been

identified into the total count of predicative adjectives.

The decisions about which tokens to include in each feature count were mainly

based on distinctions made by Hinkel (2002) and several reliable English grammars

(Biber et al., 1999; Biber et al., 2002; Quirk et al., 1985). Most features were easy to

categorize following Hinkel's system of categorization. However, certain items required

clearer operational definitions or slight modifications of categories. For example, I

counted *too* in both *too strong* and *too much strength* as an amplifier, even though one

too precedes an adjective and the other precedes a noun. While Hinkel only counted *too*

preceding an adjective, I chose to count both so that I would not get a skewed count in

case one student group tended to use the amplifier with adjectives while the other tended

to use it with nouns. In another example, Hinkel (2002) counts amplifier adverbs and

emphatics as separate categories, but after research into the differences and similarities between the two categories, I decided to count these as one feature, since several of the words on Hinkel's lists overlapped the two categories, and because their interpersonal functions are almost identical (Biber et al., 1999).

Quantitative analysis. To ensure an accurate count of each feature, I counted and re-counted the tokens in each essay. I then normed the frequency rates per 1000 words to provide a consistent basis for comparisons of texts of different lengths. The one exception, type/token ratio, was not normed per 1000. Instead, I followed the directions of Biber (1988) and counted the number of types in the first 400 words of each text and then divided this number by 400. This technique was used to prevent skewing of the type/ token ratio when comparing texts of varying lengths.

I then compared the normed frequency rates of each feature in L1 and ELL writing samples by using non-parametric statistics. A non-parametric test was appropriate because of the non-normal distribution of the data. Based on the data analysis procedures of Hinkel (2002, pp. 66-71), I used the Mann-Whitney *U* test to compare the ranks below and above the median frequency rate for each feature in each essay, between groups. I compared the medians of each data set instead of the means because a comparison of the means would result in any outliers heavily affecting the average frequency of a feature. I utilized SPSS statistical software to conduct the inferential statistics to compare the two groups' median frequency rates for each feature.

Analyzing use of features. After obtaining the statistical results of the frequency analysis, I compared concordance listings of the two corpora to find the typical discourse contexts in which each of the significantly different features were used. Using grammar reference books to identify and classify the standard functions of a feature, I looked through examples from the two corpora to determine whether the feature was used similarly in L1 and ELL writing. I also looked for examples in which the feature was used in alternate ways apart from the typical functions. I noted whether the feature was used in a semantically different way or for different discourse functions between groups, and included this information in the results and discussion section.

Answering the research questions. To answer the first research question, about features found in significantly different frequencies between the two corpora, I compared the statistical results of the Mann-Whitney U tests for each feature between groups to find the features that had significantly different median frequency rates. To answer the second research question, about the different ways that the two groups used the significantly different features, I used the analysis described above to better understand how ELL and L1 essays employed each feature in context. The following chapter presents the results of the analysis and a discussion of the findings.

Chapter 4—Results & Discussion

This chapter presents the results and discussion of the analysis. It begins with the statistical results of the Mann-Whitney U tests and a discussion of the median frequency rates and ranges of the two data sets, English Language Learner (ELL) and English First Language (L1). The second part of the chapter discusses how ELL and L1 students used the linguistic features differently in context.

Results of Frequency Analysis

Statistical results. Of the 13 features studied, the results of the Mann-Whitney U test show that five features were used in significantly different frequency rates between the two student groups: modal *would*, perfect aspect, passive voice, reduced adjective clause, and *it*-cleft (Table 4.1). These five features were used in higher frequency rates by English L1 students compared to use by ELL students. In addition, the L1 essays demonstrated a significantly higher type/token ratio than that of the ELL essays. None of the analyzed features had a significantly higher use by ELLs. The six significantly different features will be discussed in the Feature Use in Context section, with examples of how the two student groups differed in their use of these features.

The following features were not used in significantly different frequency rates: public verbs, private verbs, copula *be* as a main verb, modals of necessity, amplifiers/emphatics, predicative adjectives, and coordinating conjunctions (Table 4.2).

Table 4.1

Test Statistics for Features with Significantly Different Frequency Rates

	would	perfect	passive	adj clause	it cleft	type/token
Mann-Whitney U score	230.00	299.00	149.00	203.50	264.00	277.50
Asymp. Sig. (2-tailed)	0.001*	0.026*	0.000*	0.000*	0.000*	0.011*

*2-tailed $p \leq 0.05$.

Table 4.2

Test Statistics for Features That Did Not Have Significantly Different Frequency Rates

	public	private	be	nec-modal	amp	pred adj	coordin
Mann-Whitney U score	379.00	445.00	429.00	373.00	416.00	408.00	440.00
Asymp. Sig. (2-tailed)	0.294	0.941	0.756	0.255	0.615	0.535	0.882

Medians and ranges. For the features whose frequencies were not statistically significant, the medians were very similar between the two groups, with a difference of less than 1 word per 1000 for all but coordinators, which had a difference of 1.83 words/1000 (Table 4.3).

Except for the case of modal *would*, the ranges for frequency rates between ELL and L1 essays are similar for each feature. What's more, the ranges tend to be larger than the medians, showing a large difference between the highest and lowest frequency rates in the essays of both student groups. This demonstrates the diversity of use of the features by students within each group. The one exception, modal *would*, with a range of 7.07 for ELL and 21.15 for L1, reflects extreme outliers in the L1 essays; without these outlier frequencies, the range for modal *would* would be around 7.0, which mirrors the ELL range.

Table 4.3

Median Frequency Rates per 1000 Words (Except Type/Token Ratio) and Ranges for All Features

	ELL median	ELL range	**L1 median**	L1 range
public verbs	**7.34**	12.63	**6.99**	12.63
private verbs	**12.30**	27.45	**11.88**	23.84
be-main verb	**21.50**	24.37	**21.85**	23.67
necessity modal	**3.78**	10.04	**2.01**	8.26
modal *would*	**0.53***	7.07	**2.25***	21.15
perfect	**2.93***	12.16	**5.48***	14.77
passive	**11.51***	18.23	**18.43***	21.07
adj clause	**2.20***	8.39	**5.21***	12.52
amplifier	**7.60**	14.28	**8.30**	16.05
adj predicative	**10.31**	15.80	**9.43**	12.87
coordinator	**37.87**	36.47	**36.04**	39.02
it-cleft	**0.00***	0.93	**0.00***	2.11
type/toke ratio	**0.28***	0.15	**0.37***	0.16

*2-tailed $p \leq 0.05$.

A comparison to the findings of Hinkel (2002). All of the features that I found

to be used in higher frequencies in the L1 corpus than in the ELL corpus were consistent

with Hinkel's (2002) findings. In contrast, all of the features that were not significantly

different in the present study were found by Hinkel to be significantly more frequent in

ELL essays; Hinkel explained that they were also more common in conversation. Her

conclusion was that ELL student writing was more like spoken English and less like the

academic written register than L1 student writing. However, the results of the present

study show no evidence that ELL students use these "conversational" features more

frequently than L1 students. The median frequencies of each of these features were much

lower in the ELL writing of the present study, compared to the frequencies found in the

ELL writing of Hinkel's study.

Feature Use in Context

In this section, I present how ELL and L1 students used each of the significantly different features in the context of their essays in order to illustrate how students understood the features' functions in similar and different ways.

Modal *would*. An analysis of the usage of *would* in student writing resulted in three findings:

I.   The essays employed four different meanings of *would*, and the frequency of use for each meaning differed between L1 and ELL groups.

II.   Essay topic did not affect the frequency rate of *would* in general. However, high outlier frequencies of *would* were attributed to the essay topic.

III.   Both groups of essays contained errors*:* L1 essays employed non-standard uses of *would* and ELL essays contained ungrammatical uses of other modals in cases where *would* was obligatory.

I. <u>The meanings of *would* and their frequencies in ELL and L1 essays</u>. Hinkel describes *would* as a lexically and syntactically complex modal verb, whose "meanings are ambiguous and variable, depending on the context" (2002, p. 111). An analysis of feature use in context found 4 distinct meanings of *would* in the student essays. A closer look at the frequencies of each meaning revealed that L1 students used the various meanings of *would* in a more evenly distributed manner, while ELL students predominately used just two meanings: as a hypothetical event/state or as a politeness marker (Table 4.4). Below, I discuss each meaning in turn; refer to Table 4.4 for examples in student writing.

Table 4.4

Percentage of Use and Examples of would *by Meaning*

Meaning of *would*	L1 example	L1 %	ELL %	ELL example
1. hypothetical event or state (with or without a conditional structure)	If these slaves of human trafficking knew and believed that they were created equally …they would stand up and fight back. [310082]	48%	68%	In contrast, girls would be likely to stay if the mother were employed in either a professional or intermediate occupation. [210130]
2. a future intention expressed in the past (past tense of *will*)	However, Maryland felt that the Articles would take away many of the freedoms that the states exercised at that time and would force the states to relinquish their land claims… [310043hist]	27%	6%	…reporters were told by Mrs. Obama that she would promote breastfeeding specifically among black women, as part of her campaign to lower the children obesity. [210139]
3. a habitual or characteristic behavior, especially one that occurred in the past (Biber, et al., 1999)	Indeed, in some situations, they would even help in threatening other Muslim women who had not yet donned the robes. [320036anth]	13%	5%	When the couple knew the wife was pregnant with a baby girl, they would abandon this baby so they would have another chance to get a boy. [230131]
4. politeness or tentativeness common with requests and offers (Quirk et al., 1985)	In closing, I would like to reiterate the argument that rights and morality are simply the result of behavioral evolution. [310043ph].	1%	13%	That means, mother who tried breastfeeding for her son found perfect and significant benefits for her baby, so she would like to share it with other mothers. [210139]
Ungrammatical use or unintended/ inappropriate meaning	The advantage of having living quarters close to the porting system would allow the natives to have a first and more direct contact with those arriving in the area to conduct trade. [320087]	13%	13%	Spending time with peers changes their habits remarkably easily, and it would have some advantages and disadvantages. [210130]

1) **To convey a hypothetical event or state**: The majority of tokens in both L1 and ELL essays conveyed this meaning. However, it represented a larger percentage of tokens in ELL essays (68%) compared to L1 essays (48%).

2) **To describe a future intention expressed in the past:** This meaning conveys the past tense of modal *will.* This meaning was used in a larger percentage in L1 essays (27%) than in ELL essays (6%).

3) **To describe a habitual characteristic or behavior**: L1 essays had a higher percentage of tokens with this meaning (13%) compared to ELL essays (5%). However, it should be noted that nearly three fourths of these tokens in the L1 corpus were from one essay, which had a high outlier frequency of *would.* This essay is discussed in the sub-section about non-standard uses of *would.*

4) **To convey politeness or tentativeness**: 13% of tokens in ELL essays used this meaning, even though this meaning is not typically used in the genre of academic writing. In contrast, only 1% of L1 tokens of *would* had this meaning.

The analysis of use of each meaning shows that this feature is indeed lexically and syntactically complex, and the findings reveal a discrepancy between how ELL and L1 students employ the four meanings of *would.*

II. <u>Essay topic</u>. The topic of a given essay did not predict a higher or lower frequency of tokens of *would.* This was evidenced by the fact that although there was some overlap between L1 and ELL essays in discipline and even specific topic, group frequency rates were still significantly different. For example, both corpora include an

essay about reverse culture shock and several essays each that touch on the effects of

childhood trauma. I analyzed essay topics in both groups to check whether they

necessitated hypothetical situations (the meaning of *would* with the largest percentage of

tokens in each group) and found that even ELL essays with a zero frequency rate of

would contained *if* conditional structures co-occurring with other modals such as *will*,

might, and *should* instead of *would*:

> When a sojourner returns to his home countries, <u>if he expects</u> a happy
> return without assuming some amount of reverse culture shock, <u>he could
> have</u> a significant risk of reverse culture shock. [ELL 210144]

In most of these cases, the *if* dependent clause was used with a present tense verb and

expressed a real condition, so the use of *would* to convey a hypothetical in the

independent clause was not necessary.

In only one subset of essays did topic appear to correspond with frequency: 2 ELL

and 2 L1 essays with high outlier frequencies of modal *would*. In both ELL essays, the

topic dealt with a situation stemming from an environmental problem, and offered

hypothetical solutions to that problem:

> …if Saudi Arabia wanted to generate solar energy, <u>they would have to
> create</u> other jobs that are somewhat related to solar energy. [ELL 230153]

> Because fishermen usually sell shark fin to restaurants, reducing the
> demand from restaurants <u>would help</u> a lot in curbing shark finning. [ELL
> 230151]

Likewise, essay topic was a factor in two L1 essays with outlier frequency rates of *would*;

these essays presented hypothetical situations such as what would happen if the right to

privacy were eroded [L1 320075] and what would be the implications for the animal

Table 4.5

L1 Non-Standard Use of would *and Revised Sentences*

Examples of misuse of *would* in L1 essay	Revised sentence
[1] It was in 1843 that John Couch <u>would move</u> into the Oregon Territory and <u>would be the one</u> who was overseeing the building of the wharfs. [320087]	[2] In 1843, John Couch <u>moved</u> into the Oregon Territory and <u>oversaw</u> the building of the wharfs.
[3] …Portland has been notorious for numerous floods. As flooding <u>would occur</u>, they <u>would need</u> to clean up and clear out their storage underground. As the repairs were needed they probably <u>would decide</u> to ... [320087]	[4] …Portland <u>was</u> notorious for numerous floods. As flooding <u>occurred</u>, they <u>needed</u> to clean up and clear out their storage underground. As the repairs were needed, they probably <u>decided</u> to…

kingdom if humans had evolved differently [L1 310043ph].

      III. <u>Ungrammatical or non-standard usages</u>. The third finding was that both

groups of essays demonstrated ungrammatical or non-standard uses of *would* in the same

proportions (13% of tokens; see Table 4.4). While *would* was used more frequently in L1

essays, it appears that in some cases they over-used this device or used it in non-standard

ways. Some of these non-standard uses did not make sense in the context of the essay,

while others were standard for conversation but not for academic writing. Taken from an

essay providing a direct narrative of the history of Portland, excerpt [1] in Table 4.5

illustrates the use of *would* as a future intention expressed in the past, which is a non-

standard usage in the context of this essay. Example [1] would only be the standard usage

if at some point in the narrative, the writer explicitly took on the perspective of a

bystander during Portland's founding. Instead, the accurate way to describe a historical

fact is to use the simple past tense [2]. *Would* was also used in the same essay to convey a

past habitual behavior [3]. However, because this essay is in the discipline of history and

mainly relays past events, the typical choice is to use the simple past [4]. As the above

examples also indicate, this essay overused *would* throughout, and thus represents the L1

university student who needs more instruction on when and when not to use this modal.

Another non-standard usage of *would* found in L1 essays has a hedging effect;

this use is more typical in conversation than in academic writing. Students in both groups

used *would* in this way, perhaps because they thought it gave an 'academic tone' to their

assertions:

> There is some recent evidence showing that the norepinephrine transporter
> also normally transports some dopamine as well, <u>which would lead us to
> conclude</u> that selective norepinephrine reuptake inhibitors may also
> increase dopamine transmission as well. [L1 320092]

> This point makes this topic a more arguable topic because the social
> experience is a very important experience that <u>would most certainly effect</u>
> the academic development as well. [ELL 210145]

To illustrate how using *would* as a hedging device is appropriate for conversation, but not

for academic writing, compare the above examples to samples of spoken English taken

from the MICASE Corpus (Simpson, et al., 2002):

> There's a Miles Davis version of that which also just incredibly unique in
> its own way and <u>I would put this somewhere in that same ballpark</u>.

> <u>It seems like it would be really critical</u> to think through what the system
> wants to do and whether it's doing that.

The above comparison indicates that students are imitating conversation in their

academic writing. This tendency in L1 essays was partially responsible for the higher

frequency of *would* in the L1 corpus.

To contrast the overuse of *would* in L1 essays, the errors in ELL essays related to the underuse of *would* when students opted to use different modals where *would* was obligatory. This resulted in ungrammatical sentences:

> I always remember that my father told me he <u>will</u> always be there for me. [ELL 210157]
>
> …Students <u>must be encouraged</u> of their studying if they perceive that their parents are also cooperated with schools… [ELL 210130]

This common error corresponds to the lower frequency rate of modal *would* in ELL essays. That ELL students opt to use other modals that they understand better may indicate that they have less mastery of how to use *would* than L1 students.

The above examples of student writing reveal that ELL students are still in the process of understanding the various meanings of *would* and acquiring when to use each meaning. At the same time, L1 students have broader understanding of the various meanings of *would,* but are still in the process of acquiring how to appropriately use it in their college essays.

Perfect aspect. The lower median frequency rate of perfect aspect in ELL essays corresponds to the use of the simple past tense in cases where the perfect was required. More specifically, there were three functions of prefect aspect used in L1 essays that were often misused in ELL essays: its use for events or states starting in the past and leading up to the present time; its use with certain adverbs for making the time period explicit; and its use for presenting the findings of past research that are still considered to be relevant.

The first way in which L1 essays demonstrated the correct use of perfect aspect

was to describe an event or state that started in the past and continues up to the present

time (Biber et al., 2002, pp. 156-7):

> He and the members of his non-profit health organization… have used
> their influence to change the health policies in numerous countries. [L1
> 330078]

The meaning of the above example is the members of the non-profit organization started

to use their influence at some point in the past, and continue to use their influence today.

This contrasts with some of the ELL essays, where the simple past was used instead of

perfect aspect for this purpose:

> Smoking is one of the most reasons of illness… According to the world
> health organization (2008) an estimated 57 million people died because of
> smoking. [ELL 210158]

The writer likely meant: *…an estimated 57 million people have died because of smoking,*

as smoking began to kill people at some point in the past and continues to kill people

today. This error corresponds to the lower median frequency rate in ELL essays.

L1 essays also consistently used the perfect when the adverb accompanying the

verb required it, but ELL essays did not. These adverbs add additional information about

the preceding time leading up to the present, as demonstrated in the following L1 essay:

> They also listened to the advice of Marc Askew who has recently done
> work in Southern Thailand. [L1 310106]

In the L1 corpus, all tokens of the adverb *recently* occurred with the present perfect,

whereas in the ELL corpus, this adverb was also found with the simple past tense:

> According to Fagan, children whose <u>parents divorced recently</u> abuse drugs and alcohol much more than teenagers whose parents divorced during their early childhood. [ELL 210158]

A third way that L1 essays used perfect aspect correctly was for presenting the results of past research that are still considered valid at time of press, a common function of perfect in academic writing (Biber, et al., p.160):

> Certain frogs <u>have been found</u> to perform the amazing feat of gender flipping when the sexual ratio among the species becomes disproportionate. [L1 310043ph]

In contrast, ELL essays sometimes used simple past for this purpose, likely because they do not know about this function of perfect aspect or follow prescriptive rules to use simple past to report research findings:

> According to Lehrer (2011) <u>scientists found out that</u> old buildings make people stressed out even if were home relaxing and might be a contributing factor to having a heart attack. Some people get annoyed whenever they get in an old building. Therefore, they may not stand staying in that building and might have some serious health issues. [ELL 210143]

In the above example, the use of present tense in the second and third sentences indicates the continuing relevance of the statement made in the first sentence, *old buildings make people stressed out,* and therefore, the scientist's findings should be stated using the present perfect instead of the past tense.

Similar to the use of modal *would*, ELL essays demonstrated less mastery of when to use perfect aspect, and when not to. I have already shown examples of how the use of simple past where the perfect aspect is necessary led to ungrammaticality in ELL essays; in a final example, the use of perfect is ungrammatical or unnecessary in context:

> Parental involvement has many varieties of influence that encourage students' authentic actions…All students <u>have communicated</u> with their parents. When the children started to go to school, their parents' assistance is the most necessary to keep attending classes and enjoy the activities. [ELL 210130]

Based on the comparison of uses of perfect aspect between the two student groups, L1 essays contained correct and consistent use of perfect aspect, while ELL essays did not demonstrate mastery of this feature. Perfect aspect conveys sometimes subtle differences in meaning from simple aspect, and ELL students likely need additional practice with using perfect aspect in their academic writing.

Passive voice[1]. The lower median frequency rate in ELL essays corresponds to two prevalent errors: the violation of information structure and the ungrammatical use of another verb form where passive was required. A third factor, essay topic, did not appear to have an effect on the frequency rate of this feature.

Understanding the use of passives requires understanding typical information order in English. The principle of information structure describes the typical order of information in a clause: The subject position contains previously given information, while the predicate contains new information. Presenting information in this way makes it easier for the reader to process the information and to understand the relationships between the given information and each of the new ideas as they are presented in the text. To perform this discourse function, the subject remains in the initial position in the clause

---

[1] As described in the methods chapter, passive voice includes all passivized constructions with and without a *by*-phrase, and includes post nominal modifier –ed participle clauses, referred to as *reduced adjective clauses* in the present study.

as it changes from the agent of an action (active voice) to the recipient of an action

(passive voice):

> Pre-colonization Native Americans used more of an **indentured servant system** which, while <u>forced</u> on captives, <u>often ended</u> in integration of the captive into the tribe that they served. [No such practice] <u>was seen</u> in Anglo-American slavery where slaves were considered possessions as were their children and so on. [L1 330036]

The noun phrase in the bold font indicates the referent of the underlined verbs. After

appearing as new information in the predicate of the first clause, *indentured servant*

system stays in the subject position as the given information of subsequent clauses,

leading to the newer information being presented at the end of each clause. As new

information is presented about the *indentured servant system,* the subject goes from being

the recipient of the action (passive verb *found* in a reduced adjective clause structure)*,* to

the agent of the active verb *ended*, and then back to the recipient of *seen* (bracketed

words indicate the synonym for *indentured servant system*). This flow of information

presents the new information in the predicate position (what the reader needs to know

about *indentured servant system*) while the topic stays in subject position.

The findings of the study show that L1 essays generally used passive voice

correctly to perform the standard discourse function of this feature: ordering information

according to the principle of information structure by keeping a consistent subject as it

changes between agent of an action and recipient of an action (Biber et al., 1999).

In contrast, ELL essays tended to use active voice instead of the passive in places

where passive voice would have allowed this ordering of information. This led to the

violation of information structure, and in some cases, ungrammaticality. This violation

Table 4.6

Violation of Information Structure in ELL Student Writing

Examples from ELL essay	Revised sentence following information structure
[5] Not just <u>impact in social life</u> might affect **children** after divorce but also <u>emotional problem</u> play big role in **their** suffering. [210158]	[6] After a divorce, not only <u>are</u> **children** <u>affected</u> in their social lives, but **they** <u>may also suffer</u> from emotional problems.
As a result [of parental pressure], **some children in China** <u>are suffering</u> from depression and many emotional problems. **The education system** <u>also exposes</u> children to stress and pressure. **They** <u>rarely enjoy</u> their study process. **Children** <u>are busy preparing</u> for their exams all the time. [230131]	As a result [of parental pressure], **some children in China** <u>are suffering</u> from depression and many emotional problems. In addition, **children** <u>are exposed</u> to stress and pressure by the education system; **they** <u>rarely enjoy</u> their study process and **[ ]** <u>are busy preparing</u> for exams all the time.

makes the ideas difficult for the reader to follow. In Table 4.6, the ELL example [5] has

two different noun phrases which occupy the subject positions (underlined), while there

is a reference to *children* in the predicate of each clause (bold). The new information

emotional problem is in subject position of the second clause, which violates the principle

of information structure for creating clarity and ease of reading. A more appropriate

construction that follows this principle [6] places *children* in the subject position of both

clauses, allowing the subject to switch from a passive to active voice (underlined) as new

information about children is presented.

Further evidence of ELL students' lack of control of passive is that ELL students

used ungrammatical devices where a passive construction was the grammatical choice:

However, child abuse can also <u>be passing by</u> to the next generation. [ELL 210157].

Instead of passive voice (*child abuse can also be passed on to the next generation),* the

author uses the present progressive. The use of passive voice in this essay was well below

the median frequency rate for the ELL corpus, even though the topic of this paper, the

negative effects of having an absent father, lends itself well to use of the passive voice.

Since children are the recipients of these negative effects, one would expect a passivized

structure to be common when *child* or *children* is in the subject position. This shows that

to some extent, ELL students are still in the process of mastering when to use passive

voice, which may explain a lower frequency of use although their essay topics have just

as much potential as L1 essay topics for passive constructions. The examples below

illustrate ungrammatical constructions; in instances where the subject is ambiguous, it is

not clear whether the student chose to use active voice instead of the passive, or

attempted the passive but incorrectly left out the *be* auxiliary verb:

> They will feel much panic <u>if they involved</u> into the outside world because
> they get harmed in that world. [ELL 210159]

> That is why parents must take the step of treating their child once <u>they
> abused</u>. [ELL 210159]

Meanwhile, L1 essays employed standard academic uses of passive by

conforming to information structure and creating an effect of impersonal detachment

from a subject, conveying objectivity, and therefore accuracy, on the part of the writer

(Hinkel, 2002):

> Beta-blockers have <u>been shown</u> to retrain the sympathetic response in
> certain types of anxiety. [L1 320092]

In the above example, the writer used passive to present the statement as an objective fact

detached from the humans who conducted the experiments and discovered the results;

Table 4.7

Examples of Ineffective Choices of Voice and Revised Sentences

Examples from student writing	Revised sentences
Furthermore, spam remains untargeted: whether working or playing, domestic or overseas, no regard is shown to the users age, sex, interests, location, or computing experience. [L1 310014]	...whether working or playing, domestic or overseas, spam shows no regard to the user's age…
Absent father can affect the child many ways and **juvenile offender** is just one of the many consequences. [ELL 210157]	**Children with absent fathers** are affected in many ways, and **[ ] becoming** a juvenile offender is just one of the many consequences.

beta-blockers, the recipient of the action (they receive the actions entailed in scientific experiments), is the subject of the clause, while the agent (*researchers* or *scientists*), is omitted. The omission of an unimportant agent is another characteristic use of passive in academic writing, one which also allows for conforming to information structure (Biber et al., 2002, p. 168).

Concerning the possible effect of essay topic on frequency of use, both groups contained the same number of essays whose topics I expected would necessitate the use of passive; the fact that L1 essays have higher median frequency rate could either mean that L1 students overuse the passive voice, or that ELL students underuse this feature. Each of these possibilities is illustrated in Table 4.7, which highlights examples of where ELLs' choice of passive voice and L1s' choice of active voice led to ineffective sentences.

The examples of passive voice in student writing reveal that L1 essays keep a consistent subject to order information in accordance with the principle of information

structure, which creates an easy to follow narrative; on the other hand, ELL essays tend

to use active voice where passive would be appropriate, which leads to an ungrammatical

sentence or an ineffective or difficult to follow narrative. However, L1 essays also make

ineffective choices when choosing the passive over the active voice.

Reduced adjective clause. Two factors important for understanding ELL and L1

use of reduced adjective clauses are the choice between using a full or reduced adjective

clause and the over-generalized use of certain verbs in reduced adjective clauses, which

sometimes resulted in ungrammatical sentences.

When modifying a noun with a post-nominal participle clause, there are two

choices: a full adjective clause or a reduced form. According to a corpus analysis of

published academic writing, when participle clauses are an option, reduced adjective

clauses are more commonly used than full adjective clauses (Biber et al., 2002, p. 293).

This is likely due to the efficiency of using fewer words to convey the same meaning. In

the present study, both groups made use of full and reduced adjective clauses with both –

ed and –ing participle clauses (Table 4.8). However, the higher frequency rate of reduced

adjective clauses in the L1 corpus corresponds with a more consistent and accurate use of

this feature in L1 essays compared to ELL essays. In contrast, I found many examples of

ELL essays that used the full adjective clause where the reduced form was a more

appropriate choice (see examples in Table 4.9).

A second factor in the use of reduced adjective clauses is that ELL students

tended to use this feature in an idiomatic way with certain verbs, which often led to

Table 4.8

Examples of Full and Reduced Adjective Clauses with –ed and –ing Participle Clauses

Participle clause	Full adjective clause	Reduced adjective clause
-ed participle clause	…the Articles of Confederation, which were drafted by John Dickinson, were adopted by all of the states except Maryland. [L1 310043hi]	The first two authors listed in the citation were both students of Richard Schultes… [L1 310045]
-ing participle clause	… that composting program might help lower the cost of trash disposal in areas which were paying by weight. [ELL 210140]	…the Agri-fin Mobile program aims to provide low-income farmers as well as to benefit people living in poverty with financial services… [ELL 210132]

Table 4.9

ELL Use of Full Adjective Clauses and Revised Sentences with Reduced Adjective Clauses

Example of full adjective clause from ELL student writing	Revised sentence with reduced adjective clause
However, a survey, which was conducted by Little & Akin-Little with 149 teachers shows a result that 10% of teachers… [210138]	However, a survey conducted by Little and Akin-Little with 149 teachers shows…
The two biggest aspects that are affected by ecotourism are the environment and indigenous people. [210137]	The two biggest aspects affected by ecotourism are…

ungrammatical sentences. *Call, use,* and *need* were used at a higher percentage of tokens

of reduced adjective clauses in ELL than in L1 essays (Table 4.10). The over-generalized

use of reduced adjective clauses with these verbs led to ungrammatical uses of this

feature. The best example of this is the use of the verb *call* for identifying the name of a

concept or product when a reduced adjective clause was not necessary (Table 4.11). In

contrast, L1 essays had a higher percentage of tokens of reduced adjective clauses with

the verbs *found, given,* and *based*, and all tokens were grammatical.

Table 4.10

Verbs in Reduced Adjective Clauses with the Highest Percentage of Difference in Frequencies between ELL and L1 Corpora

Verb	ELL %	L1 %	Example from student writing
called	11.9%	3.7%	He is willing to try new products and brands from different country such as <u>a drink called Taisun's Grass Jelly Drink</u>… [ELL 210150]
used	6.3%	2.8%	…progressive relaxation is the most common <u>form of relaxation used</u> before visualization… [ELL 210133]
needed	3.2%	0.5%	…people should only pay for <u>the equipment needed</u> to generate solar energy. [ELL 210153]
found	0.8%	3.3%	…<u>cannabinods found in marijuana</u> have been known to relieve stress… [L1 320092]
given	0	3.3%	They required that <u>the aid given</u> go to industrial and agricultural development. [L1 310090]
based	0.8%	2.8%	The specialty coffee business was able to create <u>new markets</u> for itself <u>based around</u> two identity indicators… [L1 310071]

Table 4.11

Examples of ELL Essays with Unnecessary Use of Reduced Adjective Clause with Verb call *and Revised Sentences*

Example from student writing	Revised sentence
…how prepared are new couples nowadays to face the challenges of <u>the oldest institution called family</u>, which has revealed so problematic… [210158]	…how prepared are new couples nowadays to face the challenges of <u>the oldest institution, family</u>, which has revealed so problematic…
Jeff Bezos, who established Amazon.com selected <u>the name called "amazon"</u> because he hoped… [210149]	Jeff Bezos, who established Amazon.com, selected <u>the name "Amazon"</u> because he hoped…

These examples demonstrate that ELL students have more to learn about using reduced adjective clauses accurately and effectively for academic writing.

It-cleft. This feature was found rarely in both corpora, with 25 tokens found in the L1 corpus, and only 3 tokens in the ELL corpus. While the median frequencies rates for both student groups were 0.00, the ranges of frequency rates were 2.11 for L1 and 0.15

for ELL. Where students chose an *it*-cleft construction, both ELL and L1 groups used it

for the standard function; however one of two ELL students using this feature did not

form it grammatically. The low frequencies suggest that most of the students in each

group lacked knowledge about when and why to use this feature.

The discourse function of *it*-cleft is to allow the writer to focus the reader's

attention on the most important piece of information, which contributes to the cohesion of

a text (Quirk et al., 1985). The first phrase of the *it*-cleft contains the focal element of the

sentence, while the second phrase contains the background element:

> In an ironic twist of fate, **it was the US military** <u>that ushered Aristide
> back into the country</u>… [L1 330078]

In the above example, the emphasis is placed on the US Military as agent of the action—

the fact that they were responsible (bold)—instead of emphasizing the action that they

accomplished (underlined). In addition, the use of an *it*-cleft typically provides both a

connection and a clear contrast between the information provided in the *it*-cleft sentence

and the information that precedes it (Biber et al., 1999):

> <u>Culturally the men take a higher seat over women</u>, but in actuality, the
> women control a great deal of the power within the house…In the
> matrilineal society of the Minangkabau, for example, **it is the senior
> woman who keeps close track of genealogy and who ultimately decides
> the dispersal of land.** [L1 320036an]

 The information in the above *it*-cleft emphasizes that women *do* hold power in society

(bold), which contrasts with the earlier statement about how men are perceived to have a

higher status than women in this culture (underlined). In the investigation of how the

student writers used *it*-clefts, I found that each token in both student groups fulfilled the

Table 4.12

*Example from ELL Writing Where Use of i*t*-cleft Would Strengthen Emphasis and Contrast and Revised Passage*

Example from ELL essay	Revised sentence with *it*-cleft
[7] Fashion industry is highly competitive <u>but</u> the main factor to make the difference that caused Louboutin stand out is his name is not his own name. <u>On the contrary,</u> one of his signatures that can make billions of women around the world fall in love with his luxury high heels is <u>simply</u> the red sole. [210136]	[8] **It was not his name** <u>that made Louboutin stand out</u> in the highly competitive fashion industry. On the contrary, **it was the red sole,** his signature feature, <u>that made billions of women around the world fall in love</u> with his luxury high heels.

discourse functions of creating emphasis and contrast. However, the infrequent use—14

of 30 L1 essays and 2 of 30 ELL essays used *it*-clefts—suggests that the majority of the

students in each group do not have this grammatical device in their repertoire. One ELL

essay contained two tokens, but both of these *it*-clefts were ungrammatical because they

lacked the relative pronoun:

> The basic eligible populations of getting long-term care are seniors and the disabled… As we know, <u>it is not every senior and disabled person have eligibility</u>. [ELL 210134]

> Indeed, <u>it is not every poor person can be called</u> low-income people. Low-income individuals must earn less than Medicaid requires. [ELL 210134]

The above examples show that while this particular student understood the discourse

function of *it*-cleft, he or she had still not mastered how to form it correctly. Within the

scope of this study, it is not possible to determine whether ELL students did not use *it*-

clefts because they had not learned about them, or they avoided them because they did

not feel confident using them correctly.

It is important to note that students did use alternative devices to fulfill the

discourse function of contrasting two ideas such as *but, while, even though,* and *on the*

contrary. These were used with varied effectiveness, as shown in Table 4.12. The

underlined words in example [7] highlight where the student attempts to add emphasis

between contrasting ideas. However, I found that [7] was not effective in emphasizing

that Louboutin is much more than just his name, or that the red sole was his most

recognizable style. The use of an *it*-cleft adds emphasis in the way that the writer likely

intended [8].

Of all the features found to be significantly different, the *it*-cleft construction was

the least used by both groups of students, illustrating that this is not a feature that

university students are accustomed to using in their writing. It should be noted, however,

that this feature is relatively uncommon in published academic writing compared to the

other features discussed in this study; Biber et al. (2002) found the frequency of *it*-clefts

to be 600 tokens per 1 million words (equivalent to once in every four pages of an

average academic book), compared to 1400 tokens of *would,* 3500 tokens of perfect

aspect, and 20,000 tokens of passive verbs. Even the existential *there*, which is also an

alternative clause structure, was found more than four times the frequency of *it*-clefts,

with 2500 per 1 million words in academic writing. Thus, students should be cautioned to

not overuse *it*-clefts in their writing. Nevertheless, it is a useful feature with a specific

purpose in academic writing, and it will serve students well to learn how and when to use

this feature to make their writing more effective.

Type/token ratio. This feature measured the breadth of vocabulary in each essay.

As a group, ELL essays had a significantly lower type/token ratio than L1 essays, and

therefore contained less lexical diversity and lexical precision than L1 essays. It is not

Table 4.13

Number of Tokens of Three Open-Class Words and Their Percentage of Total Words in Two Essays about Reverse Culture Shock

Word	ELL essay		L1 essay	
	percentage	# of tokens	percentage	# of tokens
culture	3.12%	60	0.84%	15
shock	2.08%	40	0.56%	10
reverse	1.82%	35	0.22%	4
Total	7.02%	135	1.62%	29

surprising that English L1 students have a broader and deeper knowledge of English vocabulary than students whose English-medium education began in adulthood and who are still becoming familiar with the conventions of academic English. A comparison of vocabulary use in context revealed that L1 essays employed less repetition of vocabulary by using cohesive devices such as synonyms and demonstrative pronouns, which resulted in a higher type/token ratio for L1.

To compare the use of vocabulary between the two groups, I focused on two essays with the same topic: reverse culture shock. I compared the frequency of use of the three most common open-class words in the ELL corpus. Table 4.13 shows that the ELL use of the words *reverse, culture,* and *shock* (the three topic words) were a much higher percentage of the overall word count, compared to the use of the same words in the L1 corpus.

In the ELL essay about reverse culture shock, over 7% of total words were one of three topic words, whereas fewer than 2% of the words in the L1 essay were comprised of topic words. The reiteration of topic words throughout the essay was a common device used by ELL students to create a cohesive text:

> …lack of realistic image of returning is also one of the causes of <u>reverse culture shock</u>. If sojourners want to prevent getting <u>reverse culture shock</u>, they need to brace themselves for their return objectively. According to Martin (1984), sojourners expect different culture and some <u>culture shock</u> when they leave for a foreign country, so they can adapt easily. [ELL 210144]

The overuse of the phrases *culture shock* and *reverse culture shock* throughout the essay make the text seem repetitive and contributes to the low type/token ratio in this essay.

Meanwhile, the L1 essay has relatively fewer mentions of the topic words within the text, and instead employs other devices for creating cohesion:

> <u>Culture shock</u> is when a student has difficulty transitioning into a new culture. Occasionally, <u>these difficulties</u> prove severe enough for the student that he or she is never able to come to terms with them. [L1 310036]

Instead of using the same words to refer back to the information of the first sentence (the topic words), the writer uses the demonstrative pronoun *these* and a synonym referring to culture shock (*difficulties*) to tie together the information in the two sentences.

Besides using cohesive devices, another way that L1 students increase type/token ratio is by using synonyms throughout the essay instead of relying on the same phrase every time an idea is reintroduced:

> It is common for students to feel confused and out of place in their home country due to their personal growth, <u>changed world views</u>, and their new-found independence.
>
> After spending as much as a year with a host family, students may also begin to see the world from a <u>new perspective.</u>
>
> Students often acquire a <u>broadened point of view</u> from living and interacting with a host family and culture and want to share that with their families. [L1 310036]

The phrases *changed world views, new perspective,* and *broadened point of view* are all different ways to express the same idea. In contrast, it seems that ELL writers reuse the same words instead of diversifying their vocabulary with synonyms. The following example from the ELL essay about reverse culture shock further illustrates this tendency:

> What it comes down to is that the businessman did not <u>expect</u> any <u>difficulty</u> to fit in Israeli environment but <u>expected</u> a comfortable life. Even if the <u>sojourner</u> unconsciously <u>expects</u> a <u>happy</u> <u>return</u> but any <u>difficulty</u>, <u>sojourners</u> who are going to <u>return</u> soon should understand that they might not have the <u>happy</u> <u>return</u> which they imagine. [ELL 210144]

In the above example, with a word count of 56, there are 23 open-class words, yet 12 of these are repeats: derivations of *expect* and *return* have 3 tokens each, while *difficulty, happy,* and *sojourner(s)* each have 2 tokens. The total number of unique open-class words in this passage of 56 words is actually only 16. The repetition of words instead of the use of cohesive devices and synonyms correspond to the significantly lower type/token ratio in ELL essays.

Summary

The first research question asked whether or not the writing of ELL students in the IELP and English L1 students at Portland State University contained significantly different frequency rates of the linguistic features that Hinkel (2002) found to be different between the two student groups. In this chapter I presented the results of the Mann-Whitney U tests, which answered the research question: There were in fact six linguistic features found in significantly lower frequencies in ELL essays than in English L1 essays. To answer the second research question, which sought to understand the

differences in how students used each of the six significantly different features in their writing, I analyzed examples of each feature within ELL and L1 essays and compared how each group used the feature in the context of the essays. In the next chapter I will summarize the most important findings and discuss their implications for teachers and students.

Chapter 5— Conclusion

The final chapter provides a summary of the findings presented in the results and discussion chapter, followed by the implications and applications of the findings to the teaching of academic writing to English language learners. The chapter concludes with a discussion of the limitations of the study and suggestions for future research.

Summary of the Results

In order to better understand the preparedness of students in the Intensive English Language Program to take undergraduate courses at Portland State University, this study sought to find differences between the frequencies and uses of 13 linguistic features in essays written by ELL students in the final level of academic English instruction and English L1 students in mainstream undergraduate courses. The analysis of the frequencies of linguistics features in L1 and ELL essays and their use in the context of student writing resulted in several interesting findings:

- Seven features were not found in significantly different frequencies: public verbs, private verbs, *be* as main verb, modals of necessity, amplifiers, predicative adjectives, and coordinating conjunctions.
- These features are associated with conversation and were the features that Hinkel (2002) found in higher frequencies in ELL writing. That this study did not find these features in significantly different frequencies between L1 and ELL essays means that ELL students in this study did not rely more heavily than L1 students on the 7 features typical of conversation.

- Six features were found in significantly different frequencies between the two groups: modal *would,* perfect aspect, passive voice, reduced adjective clauses, *it-*clefts, and type/token ratio.

- All of the features found to be significantly different were more frequent in L1 essays. These features are associated with published academic writing. For the most part, L1 students demonstrated mastery of these in their writing, while ELL essays contained ungrammatical usages of these features, or omitted them where they were obligatory. From these observations I concluded that ELL students were still in the process of learning how to use these features in their writing.

- *Would* was found in significantly lower frequencies in ELL essays than in L1 essays. ELL essays predominantly used two meanings of *would*, one of which, to convey politeness, is not a typical use in academic writing. Meanwhile, L1 essays used the four meanings of *would* in a more evenly-distributed manner. However, both groups of students used *would* in ungrammatical or non-standard ways in the same proportions. This shows that ELL students are still in the process of understanding the various meanings of *would* and acquiring when to use *would*, while both groups of students are still learning how to use this feature in their academic writing.

- Perfect aspect, passive voice, and reduced adjective clauses were used in significantly lower frequencies in ELL essays than in L1 essays. Based on an analysis of how each group used these features in context, it appeared that L1 essays used these features in accordance with the standard uses in academic

writing. In contrast, ELL essays tended to use these features in ungrammatical or non-standard ways, and at the same time they used other features where perfect, passive, or reduced adjective clauses were obligatory or were stylistically a more appropriate choice.

- *It*-clefts were used in ELL essays with significantly lower frequency than in L1 essays, although they were used infrequently by both groups. While some students in both groups used this feature correctly for its standard purpose, the majority of students appeared to not have this feature in their repertoire. Instead, they used other devices for emphasizing a contrast.

- As a group, ELL essays had a lower type/token ratio than L1 essays. This corresponded with ELL essays repeating topic words to refer back to ideas in a text. In contrast, L1 essays used a greater breadth of vocabulary by utilizing demonstrative pronouns and synonyms instead of repetition to create cohesion, which corresponded to a higher type/token ratio in L1 essays.

While teachers of English-medium content courses may tend to see ELL papers as "weaker" than those written by L1 students, this weakness cannot be entirely explained by the lower frequency of use in ELL writing of the six features mentioned above. Based on the analysis of use of features in context, it seems unlikely that the underuse of these features alone have a global impact on the ELL essays, nor does it account for teacher perceptions that ELL papers are less effective; what affected the quality of the essay more noticeably was when ELL student used one of these features in an ungrammatical way, or in cases where another feature would have been more effective. In addition, rhetorical

features are likely responsible for teacher perceptions of writing quality and effectiveness. An analysis of rhetorical features was beyond the scope of this study, but this is an area of future research and will be discuss later in the chapter.

An additional interpretation of the results addresses students' understanding of academic writing. The results show that ELL students are still in the process of acquiring certain features to use in their writing, namely modal *would*, perfect aspect, passive voice, reduced adjective clause, and it-clefts. While ELL students do use these features, they use them in noticeably different ways compared to the L1 students. ELL essays also had a significantly lower type-token ratio which means they must continue to develop their vocabulary and broaden their understanding of cohesive devices in academic writing to produce a higher quality of writing.

The results also offer implications for L1 student writing: Based on their essays, L1 students seemed to have acquired the uses of perfect aspect, passive voice, and reduced adjective clauses appropriate for academic writing. However, their use of *would* demonstrated a lack of appreciation of register, sometimes resembling usages typical of conversation instead of those typical of academic writing. In addition, they used *it*-clefts very infrequently, implying that they might need further instruction to better understand the value of employing this feature in their writing.

In comparing the findings of the present study with those of Hinkel (2002), on which the present study was based, both similarities and differences emerge. Both studies found significantly higher frequency rates of the following features in L1 essays: *would,*

perfect aspect, passive voice, reduced adjective clauses, and *it*-clefts (type/token ratio was

not analyzed in Hinkel, 2002). That the ELL essays in this study contained significantly

lower frequencies of these features corroborates Hinkel's claim that ELL students have a

narrower grammatical range than L1 students in their academic writing. However,

because Hinkel did not go into much detail describing the context in which the two

groups of students used the significant features, the present study contributes a deeper

understanding of how ELL students use features typical of academic writing, and how

closely these usages align with those in L1 student writing. These comparisons provide

implications for how instructors can help ELL students further build their writing skills as

students work towards becoming proficient in academic English.

　　　The results of the current study differ from Hinkel's in that they did not reveal

significant differences in the frequencies of features associated with conversation, which

Hinkel found in higher frequencies in ELL writing (public and private verbs, *be* as a main

verb, necessity modals, amplifiers, predicative adjectives, and coordinating

conjunctions). It is important to note that the L1 and ELL essays in the present study

contained lower frequency rates of these features than did the essays in Hinkel's study,

meaning that PSU students demonstrated less use of conversational features in their

academic writing than did the students in Hinkel's study. Some variability in results

could be due to the six different L1 groups represented in Hinkel's study. The present

study treated ELL writers as one group, not accounting for differences in features use

between specific L1 groups. The differing results may also be explained by differences

between the samples of student writing used in the respective studies. The samples of L1

and ELL writing used in Hinkel's study were timed argumentative essays written as

diagnostic tests; the high-pressure environment of a timed writing test requires the real-

time mental processing associated with conversation, which may explain the

conversational features found in ELL writing samples. In contrast, the writing samples in

the present study were research essays that had been revised through multiple drafts and

had presumably taken several weeks to write. As demonstrated by the similar frequencies

of the aforementioned features in ELL and L1 writing, ELL students were likely alerted

to any existing features of conversation in their essays during the writing process, and

were given the chance to change them in accordance with appropriate academic

language. Compared to the writing samples analyzed in Hinkel's study, the writing

samples I analyzed were more representative of the writing that ELL students will be

expected to produce in mainstream university courses: research essays written in multiple

drafts. This is the type of writing taught in the specific context of Portland State's IELP

to prepare ELL students for university coursework. In the next section, I discuss how the

findings of this study could be applied to teaching academic writing.

Applications for Language Teaching

Because the purpose of this study is to learn more about the writing ability of

students in Portland State University's IELP, it is my hope that the findings of this study

will be relevant first and foremost to the instructors in the IELP, who will be able to refer

to a statistical analysis of student writing for a better understanding about the features that

ELL students underuse in their writing assignments. By shedding light on the gaps that

exist between English L1 knowledge of academic writing and that of IELP students, the

findings may be used by the IELP instructors and administrators to inform the content of

the level 5 writing course, or influence the creation of materials and future textbooks for

IELP writing courses. The student workers in the IELP Learning Center, who tutor IELP

students in their academic work, will similarly be able to use the findings of this study to

target certain features when a student come to them asking for help on their term paper.

Below I offer some specific applications for teachers and tutors.

Modal *would*. The examples of student writing reveal that although L1 students

seemed to have more knowledge than ELL students about how to use the various

meanings of *would*, neither group demonstrated mastery of this feature in their academic

writing. A tendency to predominantly rely on only two meanings of *would,* and many

ungrammatical tokens of this feature show that ELL students are still in the process of

acquiring when to use *would* and understanding its various meanings in general. At the

same time, overuse and non-standard use of this feature in L1 essays show that L1

students are still acquiring how to convey precise meanings using *would* in their college

essays. Therefore, both L1 and ELL college students would likely benefit from being

taught explicitly in their composition classes how and when to use this modal

appropriately in academic writing, and when this device is unnecessary or only

appropriate for conversation. For example, L1 students used *would* to hedge when

making a claim in their academic writing (e.g. "I would say…"), while this use of *would*

is more common to conversation. Instructors of academic writing for ELL and L1

university students should do activities that bring this feature into students' awareness,

and that train students to distinguish between the uses of *would* that are appropriate for

academic writing and those that should only be used in conversation. The following

sample activity helps students to notice, discuss and use *would* in academic text:

1. Students read a series of passages and determine whether each passage is from

conversation or academic writing and explain their answer.

2. Students circle all tokens of *would* in each passage.

3. Students identify the time frame and speaker identity in each passage, and the purpose

for using *would* in the context of the passage.

4. The instructor teaches students the various meanings of *would*.

5. Students match these meanings to the examples of *would* that they identified.

6. Students and teacher talk about which meanings of *would* are appropriate for use in

academic writing, taking examples from the sample passages.

7. Students look at another handout of ungrammatical and uncommon uses of *would*, and

collaborate on fixing each passage to follow the standards of academic writing.

This sample activity is illustrated by two student worksheets on the following pages

(Figures 5.1 and 5.2).

Figure 5.1: Sample materials: Contrasting uses of modal *would* in conversation and academic writing

---

 MODAL *would* ACTIVITY I:

1. For each of the passages, determine whether it comes from conversation (C) or academic writing (AW). Write C or AW on the line and explain your answer.

Passage A:

> **Yeah, I mean, I talked to the restaurant waitress that would wait on them Saturday mornings, and she was a-gasp that we could have charged John Hamilton, because she would see them at breakfast together all lovey-dovey.**

> Register:_____ How do you know?

Passage B:

> **In January, the Defense Department met with scholars, who left the meeting convinced that their concerns had been heard and that the museums and ancient sites of Iraq would be protected.**

> Register:_____ How do you know?

Passage C:

> **Look, I – I just want you protected. That's all I know. Please, would you ask Chris to come in? I really need to talk to him.**

> Register: _____ How do you know?

Passage D:

> **For instance, a mechanic wearing a haptic glove could simulate reaching into a tight space and turning a screw. If the space were too narrowly designed for a human hand to perform the motion easily, the engineer would know to redesign that part of the line without ever having built a prototype.**

> Register: _____ How do you know?

Figure 5.1 (continued): Sample materials: Contrasting uses of modal *would* in conversation and academic writing

2. Circle *would* in each passage. Then fill out the chart for the time frame and speakers of each passage, and purpose that *would* has in each passage.

Passage	**Time frame**: tense and aspect? Habitual or completed action?	**Speakers**: Pronouns (1st, 2nd, 3rd person)? Clues about identity?	**Purpose**: Why does the speaker/writer use would?
A			
B			
C			
D			

3. Read each passage again. Then determine the register that uses each meaning, and the passage that contains an example of the meaning.

	Meaning:	Register:	Passage:
a.	Hypothetical event or state	_____	_____
b.	Future in the past tense	_____	_____
c.	Habitual past action	_____	_____
d.	Politeness/tentativeness	_____	_____

4. Complete the sentences:

   a. In academic writing *would* is usually used to express _____ and _____.

   b. In conversation, *would* is used to express the two meanings above, plus the other two meanings: _____ and _____.

Figure 5.2: Sample materials: Correcting uncommon uses of *would* in academic writing

<u>MODAL *would* ACTIVITY II:</u>

In the examples of academic writing below, circle each example of *would.* Then fill in the chart to answer the questions for each passage.

1. It was in 1843 that John Couch would move into the Oregon Territory and would be the one who was overseeing the building of the wharfs.

2. In closing, I would like to reiterate the argument that rights and morality are simply the result of behavioral evolution.

3. Portland has been notorious for numerous floods. As flooding would occur, they would need to clean up and clear out their storage underground.

	Passage 1	Passage 2	Passage 3
Which meaning of *would* is used-- **future in the past, habitual past**, or **politeness**?			
Is this use common in academic writing? (refer to Activity I)			
Paraphrase the passage without using *would.*			

4. Based on your answers above, how do you convey each purpose in academic writing?

Purpose	Conversation	Academic writing
A past event that has finished	simple past	
Politeness	*would like*	
Repeated past events	*would* + base form of verb	

Perfect aspect, passive voice, and reduced adjective clauses. L1 students seem to have already mastered these linguistic features, while ELL students are still in the process of acquiring these features in their academic writing. Because the ELL students have completed through level 4 of intensive English grammar courses, it is assumed that they have been taught about these features possibly a number of times. However, the instruction may not have included enough practice producing the various uses of each of these features in academic text.

In the case of reduced adjective clauses, ELL students may default to using a full adjective clause in post-nominal position instead of the reduced adjective clause; they have had more instruction and practice in how to form and when to use full adjective clauses, since these are introduced before the reduced adjective clauses in grammar books. What's more, reduced adjective clauses are very difficult for English language learners because sentences appear to have two main verbs. On the other hand, English L1 students at the university level have had more exposure than ELL students to how academic writing 'sounds', and while they might not have been taught about participle clauses, the use of this feature may be built into their native intuition about what 'sounds right'. Not only should teachers remind students that there is a choice between the full and reduced form, but also point out cases where an adjective clause should not be used at all.

Instructors of academic writing for ELL students should regularly engage students in practicing the use of perfect, passive, and reduced adjective clauses in their own writing. This begins with noticing usages of the features and being able to identify what

makes these usages appropriate and effective. Also, students should practice noticing and

correcting the absence of these features in cases where they are obligatory. Instructors

can help students reinforce knowledge of these features and improve their ability to use

them in writing through doing noticing activities similar to the one introduced for

teaching modal *would.*

It-cleft. Of all the features found to be significantly different, the *it*-cleft

construction was the least used by both groups of students. While the frequencies seem to

suggest that L1 student have more experience with cleft constructions and understand

how to form them and when to use them, it is likely that L1 students who are still

obtaining their undergraduate education underuse this feature compared to writers at the

professional level. Unlike the ELL students, who have received intensive English

grammar instruction over many terms, the L1 students likely have not received explicit

instruction about the form and function of cleft constructions. If the form and function of

it-cleft were taught in freshman composition courses, students may end up utilizing this

construction more often in their writing.

The *it*-cleft is an effective device for emphasizing contrast, which is especially

useful for argumentation in academic writing. Instructors of both ELL and L1 students

should encourage students to utilize this feature in their writing by teaching them its

specific purpose and giving them opportunity to practice using it in an academic text. It

may also be helpful to show students examples of *it*-clefts in published writing and then

have students practice adding this structure to sample sentences which lack the

appropriate emphasis. However, instructors should caution students to not over-use this

feature; to use it effectively, students have to be able to identify the appropriate context in which to use it.

       Type/token ratio. Based on the observation that ELL students tend to over-rely on the repetition of key words to refer back to ideas in the text, it seems that ELL students do not know how to create cohesion in a text without repeating words. It seems that are not yet adept at how to utilize cohesive devices for this purpose. In addition, it is possible that students focus their attention on more significant issues such as grammar and content, so using a broader vocabulary to avoid repetition is an afterthought. Writing instructors should encourage students to develop their academic vocabulary, emphasizing that a wide range of vocabulary strengthens academic text by increasing the precision of their ideas.

Limitations and Suggestions for Future Research

       There are several limitations to my study. First, I analyzed two small corpora of thirty papers each; with a larger corpus of student writing, patterns may have emerged that were not great enough to be significant in the corpora I compiled. Second, I analyzed thirteen linguistic features in each text, while some published studies have analyzed over sixty features (Hinkel, 2002; Biber, 1988). This means that there may be features that I chose not to test that actually do occur in significantly different frequencies between the two groups of essays. Third, while I did not differentiate ELL essays by the writer's L1, the comparison of ELL writing between L1 groups might have revealed particular disparities or needs for intervention that were not made evident by grouping all ELL

essays together. While the scope of a master's thesis made the study of a more limited

number of texts and features appropriate, future research should include a larger corpus

of student writing, a wider selection of linguistic features, and a comparison of feature

use by L1 group to lend deeper insight into ELL student writing.

In this study, the analysis of feature use in context was limited to the features that

were found in significantly different frequencies. Even though the features associated

with conversation were found in similar frequency rates between the two groups, further

research could reveal how ELL and L1 essays use conversational features in their writing,

and whether they use them for similar or different functions.

One limitation is also a strength of the study: All of the texts were written by

students at one university. On one hand, this means that the findings are less likely to be

applicable to students at other universities or in different levels of intensive English

programs. On the other hand, this design is integral to the purpose of my study, in that the

findings are intended to bring about a better understanding of the writing of a specific

group of students, the international students studying in the IELP, and not the writing of

students in a different intensive English program. In the future, graduate students or

professional teachers at other universities could use this study as a model for carrying out

research on student writing in their respective English language programs.

An additional area of future research is to delve deeper into the differences

between ELL and L1 academic writing by performing a rhetorical analysis in addition to

linguistic features analysis to determine if there is correlation between the frequency of

each linguistic feature and the essay's score based on a teacher's observations. Another possible rhetorical analysis could be a quantitative study examining idea development, idea organization, integration of ideas to see if there a significant difference between the organizational features and discourse structure of the two groups.

In addition to a quantitative analysis, a related qualitative study would connect the frequency of linguistic features with teacher perceptions of what qualifies as "good" writing. Future researchers could interview professors in various disciplines about their expectations of student writing and ask them to identify examples of what they determine to be good student writing. The survey of professors and the linguistic analysis of student writing would be woven together to illuminate which features students should include in their writing. Another possible qualitative study would shed light on the students' perspective of what "good writing" is, the techniques they use to achieve good writing, what features they think are important in academic writing, and what messages they have received throughout their education about writing well.

Lastly, an analysis of formulaic language in the two groups of essays and in published academic writing would lend insight into the common collocations used by L1 students and professionals, but not by ELL students. This would inform writing instructors of the specific examples of formulaic language that would be helpful for ELL students to use in their own academic text.

Conclusion

This study contributed to the field of English language teaching, specifically the instruction of academic writing for ELL students, in that it identified a gap between the writing of ELL students and the L1 students with whom they (will) take university courses. This gap points to recommendations for the features that writing instructors should reinforce in advanced ELL writing courses. Through continuing efforts to better understand our ELL students' writing capabilities and shortfalls, teachers of English for Academic Purposes can help their students reach their highest potential for success in U.S. universities.

References

Aijmer, K. (2002). Modality in advanced Swedish learners' written interlanguage. In S. Granger, J. Hung, and S. Petch-Tyson (Eds.), *Computer learner corpora, second language acquisition and foreign language teaching* (pp. 55-76). Philadelphia: John Benjamins Publishing.

Albers, S. (2007). *The Viking Corpus of Student Academic Writing* (Unpublished master's project). Portland State University, Portland, OR.

Altenberg, B., & Tapper, M. (1998). The use of adverbial connectors in advanced Swedish learners' written English. In S. Granger (Ed.), *Learner English on computer* (pp. 80-93). New York: Addison Wesley Longman Inc.

Barlow, M. (2004). MonoConc Pro 2.0. Houston, TX: Athelstan Publications.

Beck, J. H. (2007). *The register of academic writing*: *A multidimensional comparison of German advanced EFL leaners and American native speakers* (Unpublished master's thesis). Portland State University, Portland, OR.

Benesch, S. (1993). ESL, ideology, and the politics of pragmatism. *TESOL Quarterly, 27,* 705-717.

Benesch, S. (2001). *Critical English for academic purposes.* Mahwah, NJ: Lawrence Erlbaum Associates.

Biber, D. (1988). *Variation across speech and writing.* New York: Cambridge University Press.

Biber, D. (2006). *University language*. Philadelphia: John Benjamins Publishing.

Biber, D., & Conrad, S. (2009). *Register, genre, and style.* New York: Cambridge University Press.

Biber, D., Conrad, S., & Leech, G. (2002). *Longman student grammar of spoken and written English.* Essex: Pearson Education Limited.

Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: investigating language structure and use.* New York: Cambridge University Press.

Biber, D., Johansson, S., Leech, G., Conrad, S., Finegan, E. (1999). *Longman grammar of spoken and written English.* Harlow, England: Longman.

Biber, D., & Reppen, R. (1998). Comparing native and learner perspectives on English grammar: a study of complement clauses. In S. Granger (Ed.), *Learner English on computer* (pp. 145-159). New York: Addison Wesley Longman Inc.

Bitchener, J., & Knoch, U. (2010). Raising the linguistic accuracy level of advanced L2 writers with written corrective feedback. *Journal of Second Language Writing, 19,* 207–217. doi:10.1016/j.jslw.2010.10.002

Breyer, Y. A. (2011). *Corpora in language teaching and learning.* Frankfurt: Peter Lang.

Casanave, C. P. (2004). *Controversies in second language writing*. Ann Arbor: University of Michigan Press.

Chang, C-F., & Kuo, C-H. (2011). A corpus-based approach to online materials development for writing research articles. *English for Specific Purposes, 30,* 222-234.

Chapelle, C. (1998). Some notes on systemic-functional linguistics. *International Systemic Functional Linguistics Association.* Retrieved from http://www.isfla.org/Systemics/documents/chapelle.html

Chen, Y-H. (2010). Lexical bundles in L1 and L2 academic writing. *Language Learning & Technology, 14,* 30-49.

Chen, Y-S., & Su, S-W. (2012). A genre-based approach to teaching EFL summary writing. *ELT Journal: English Language Teachers Journal, 66,* 184-192.

Cheng, F-W. (2008). Scaffolding language, scaffolding writing: a genre approach to teaching narrative writing. *Asian EFL Journal, 10,* 167-191.

Conrad, S. (1996). Investigating academic texts with corpus-based techniques: An example from biology. *Linguistics and Education, 8,* 299-326.

Crossley, S. A., & McNamara, D. S. (2009). Computational assessment of lexical differences in L1 and L2 writing. *Journal of Second Language Writing, 18,* 119-135.

Crossley, S. A., & McNamara, D. S. (2012). Predicting second language writing proficiency: The roles of cohesion and linguistic sophistication. *Journal of Research in Reading, 35,* 115–135. doi:10.1111/j.1467-9817.2010.01449.x

Davies, M. (2008-). *The Corpus of Contemporary American English: 450 million words, 1990- present*. Available online at http://corpus.byu.edu/coca/.

Diab, N. M. (2010). Effects of peer- versus self-editing on students' revision of language errors in revised drafts. *System, 38,* 85–95. doi:10.1016/j.system.2009.12.008

Evans, N. W., Hartshorn, K. J., & Strong-Krause, D. (2011). The efficacy of dynamic written corrective feedback for university-matriculated ESL learners. *System*, *39,* 229–239. doi:10.1016/j.system.2011.04.012

Gardner, S. (2012). Genres and registers of student report writing: An SFL perspective on texts and practices. *Journal of English for Academic Purposes, 11,* 52-63.

Gillard, P., & Gadsby, A. (1998). Using a learners' corpus in compiling ELT dictionaries. In S. Granger (Ed.), *Learner English on computer* (pp. 159-171). New York: Addison Wesley Longman Inc.

Gilquin, G., Granger, S., & Paquot, M. (2007). Learner corpora: The missing link in EAP pedagogy. *Journal of English for Academic Purposes, 6,* 319-335.

Granger, S. (1998). The computer learner corpus: a versatile new source of data for SLA research. In S. Granger (Ed.), *Learner English on computer* (pp. 3-18). New York: Addison Wesley Longman Inc.

Granger, S. (2003). The International Corpus of Learner English: A new resource for foreign language learning and teaching and second language acquisition research. *TESOL Quarterly, 37,* 538-546.

Hinkel, E. (2001). Matters of cohesion in L2 academic texts. *Applied Language Learning, 12,* 111-132.

Hinkel, E. (2002). *Second language writers' text: Linguistic and rhetorical features.* Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Hinkel, E. (2003a). Simplicity without elegance: Features of sentences in L1 and L2 academic texts. *TESOL Quarterly, 37,* 275-301.

Hinkel, E. (2003b). Adverbial markers and tone in L1 and L2 students' writing. *Journal of Pragmatics, 35,* 1049-1068.

Hinkel, E. (2004). Tense, aspect and the passive voice in L1 and L2 academic texts. *Language Teaching Research, 8,* 5-29.

Hinkel, E. (2009). The effects of essay topics on modal verb uses in L1 and L2 academic writing. *Journal of Pragmatics, 41,* 667-683.

Hopkins, K. (2011, December 20). 10 U.S. colleges with highest ESL participation rates. *U.S. News and World Report.* Retrieved on April18, 2013 from http://www.usnews.com/

Hyland, K. (2000). *Disciplinary discourses: social interactions in academic writing.* New York: Longman.

Hyland, K., & Milton, J. (1997). Qualification and certainty in L1 and L2 students' writing. *Journal of Second Language Writing, 6,* 183-205.

Institute of International Education. (1955- 2011). *Open doors: Fast facts international students in the U.S.* Washingon, DC/ New York: Author.

Intensive English Language Program. (n.d.). Retrieved from http://www.ielp.pdx.edu/programs.php

IELP. (2013). *Program Review Data.* [Graphs and spreadsheets of student demographic and enrollment information]. Portland State University Intensive English Language Program, Portland, OR.

Jalilifar, A. (2010). Thematization in EFL students' composition writing and its relation to academic experience. *RELC Journal, 41,* 31–45. doi:10.1177/0033688210362621

Kathpalia, S. S., & Carmel, H. L. H. (2011). Metaphorical competence in ESL student writing. *RELC Journal, 42,* 273–290. doi:10.1177/0033688211419379

Kongpetch, S. (2006). Using a genre-based approach to teach writing to Thai students: A case study. *Prospect: An Australian Journal of Teaching/Teachers of English to Speakers of Other Languages (TESOL), 21,* 3-33.

Kormos, J. (2011). Task complexity and linguistic and discourse features of narrative writing performance. *Journal of Second Language Writing, 20,* 148-161. doi:10.1016/j.jslw.2011.02.001

Laufer, B., & Waldman, T. (2011). Verb-noun collocations in second language writing: a corpus analysis of learners' English. *Language Learning, 61,* 647-672.

Lorenz, G. (1998). Overstatement in advanced learners' writing: stylistic aspects of adjective intensification. In S. Granger (Ed.), *Learner English on computer* (pp. 53-66). New York: Addison Wesley Longman Inc.

Martin, J. R., & Rose, D. (2008). *Genre relations: mapping culture.* London: Equinox.

Maseeh College of Engineering and Computer Science. (2011). Maseeh College of Engineering and Computer Science Arts & Letters and Social Science Requirement. Retrieved from http://www.pdx.edu/sites/ www.pdx.edu.cecs/files/AL-SS_Approved%205_20_2011.pdf

Negretti, R., & Kuteeva, M. (2011). Fostering metacognitive genre awareness in L2 academic reading and writing: A case study of pre-service English teachers. *Journal of Second Language Writing, 20,* 95–110. doi:10.1016/j.jslw.2011.02.002

NEW University Writing Requirements. (n.d.). Retrieved from http://www.pdx.edu/advising/writing-requirement

Nurmukhamedov, U., & Kim, S. H. (2010). 'Would you perhaps consider…': Hedged comments in ESL writing. *ELT J, 64,* 272–282. doi:10.1093/elt/ccp063

Ong, J. (2014). How do planning time and task conditions affect metacognitive processes of L2 writers? *Journal of Second Language Writing, 23,* 17-30. doi:10.1016/j.jslw.2013.10.002

Ortega, L. (2003). Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing. *Applied Linguistics, 24,* 492-518. doi:10.1093/applin/24.4.492

Portland State University Facts. (n.d.). Retrieved from http://www.pdx.edu/portland-state-university-facts

Profile. (n.d.). Retrieved from http://www.pdx.edu/profile/snapshot-portland-state

Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A comprehensive grammar of the English language.* London: Longman.

Rafiei, K., & Modirkhamene, S. (2012). Thematicity in published vs. unpublished Iranian TEFL theses. *Theory & Practice in Language Studies, 2,* 1206-1213.

Ramanathan, V., & Atkinson, D. (2011). Individualism, academic writing, and ESL writers. In P. K. Matsuda, M. Cox, J. Jordan, & C. Ortmeier-Hooper (Eds.), *Second language writing in the composition classroom* (pp. 165-191). Boston: Bedford/ St. Martin's.

Reppen, R. (2012). *Grammar and beyond level 1 student's book.* New York: Cambridge University Press.

Santos, T. (1992). Ideology in composition: L1 and ESL. *Journal of Second Language Writing, 1,* 1-15.

Santos, T. (2001). The place of politics in second language writing. In T. J. Silva and P. K. Matsuda (Eds.), *On second language writing* (pp. 173-190). Mahwah, N.J.: L. Erlbaum Associates.

Shaw, P. (2009). Linking adverbials in student and professional writing in literary studies: What makes writing mature. In M. Charles, D. Pecorari, & S. Hunston (Eds.), *Academic writing: At the interface of corpus and discourse* (pp. 215-235). New York: Continuum.

Simpson, R. C., Briggs, S. L., Ovens, J., & Swales, J. M. (2002). *The Michigan Corpus of Academic Spoken English.* Ann Arbor, MI: The Regents of the University of Michigan.

Siyanova, A., & Schmitt, N. (2008). L2 learner production and processing of collocation: A multi-study perspective. *The Canadian Modern Language Review, 64,* 429-58.

Smith, D. (2010.) IELP Grammar and Writing Curriculum. Updated by M. Mulder, 2012. Portland State University, Portland, OR

Smith-Barrow, D. (2013, December 24). 10 Colleges with the highest percentage of students in ESL. *U.S. News and World Report.* Retrieved on January 31, 2014 from www.usnews.com.

Swales, J. M. (1990). *Genre analysis: English in academic and research settings.* New York: Cambridge University Press.

Terzian, S. G., & Osborne, L. A. (2006). Postwar era precedents and the ambivalent quest for international students at the University of Florida. *Journal of Studies in International Education*, *10,* 286-306. doi: 10.1177/1028315306286919

University Studies Goals. (n.d.). Retrieved from http://www.pdx.edu/unst/university-studies-goals

UNST Introduction. (n.d.). Retrieved from http://www.pdx.edu/unst/unst-introduction

Wei, J. (2013). Corpus-based research on topical theme choices in Chinese and Swedish English learner writings. *Theory & Practice in Language Studies, 3,* 2202-2208.

Yasuda, S. (2011). Genre-based tasks in foreign language writing: developing writers' genre awareness, linguistic knowledge, and writing competence. *Journal of Second Language Writing, 20,* 111–133. doi:10.1016/j.jslw.2011.03.001

Yayli, D. (2011). From genre awareness to cross-genre awareness: A study in an EFL context. *Journal of English for Academic Purposes, 10,* 121–129. doi:10.1016/j.jeap.2011.02.001