REPUBLIC OF TURKEY
ÇUKUROVA UNIVERSITY
THE INSTITUTE OF SOCIAL SCIENCES
DEPARTMENT OF ENGLISH LANGUAGE TEACHING

STANCE LEXICAL BUNDLES IN ACADEMIC L2 WRITING: A CONTRASTIVE INTERLANGUAGE ANALYSIS

Meltem MUŞLU

A PhD DISSERTATION

ADANA / 2014

REPUBLIC OF TURKEY
ÇUKUROVA UNIVERSITY
THE INSTITUTE OF SOCIAL SCIENCES
DEPARTMENT OF ENGLISH LANGUAGE TEACHING

STANCE LEXICAL BUNDLES IN ACADEMIC L2 WRITING: A CONTRASTIVE INTERLANGUAGE ANALYSIS

Meltem MUŞLU

Supervisor: Assoc. Prof. Dr. Cem CAN

A PhD DISSERTATION

ADANA / 2014

To the Directorship of the Institute of Social Sciences, Çukurova University

We certify that this dissertation is satisfactory for the award of the Degree of Doctor of Philosophy.

Supervisor: Assoc. Prof. Dr. Cem CAN

Member of Examining Committee: Prof. Dr. Hatice SOFU

Member of Examining Committee: Assoc. Prof. Dr. Ahmet DOĞANAY

Member of Examining Committee: Assoc. Prof. Dr. Türkay BULUT

Member of Examining Committee: Assoc. Prof. Dr. Jülide İNÖZÜ

I certify that this dissertation confirms to the formal standards of the Institute of Social Sciences.

…../…../2014

Prof. Dr. Yıldırım Beyazıt ÖNAL

Director of the Institute

Note: The uncited usage of the reports, charts, figures, photographs in this thesis, whether original or quoted from other sources, is subject to the law of Works of Art and Thought No: 5846.
Not: Bu tezde kullanılan özgün ve başka kaynaktan yapılan bildirişlerin, çizelge, şekil ve fotoğrafların kaynak gösterilmeden kullanımı, 5846 Sayılı Fikir ve Sanat Eserleri Kanunu'ndaki hükümlere tabidir.

ÖZET

İKİNCİ DİL AKADEMİK YAZIMLARINDAKİ TUTUM SÖZCÜK ÖRÜNTÜLERİ: BİR KARŞILAŞTIRMALI ARADİL ANALİZİ

Meltem MUŞLU

Doktora tezi, İngiliz Dili Eğitimi
Danışman :Doç. Dr. Cem CAN
Aralık, 2014, 92 sayfa

Derlem dilbilim ‘derlem olarak bilinen ve original metinlerin elektronik ortamlarda kaydedilmesi ve kullanılması üzerine kurulu olan bir dilbilim yöntemidir (Granger, 2002, s. 4). Son 50 yılda bilgisayar teknolojisinde elektronik verilerin uygulama hızlarının ve saklama kapasitelerinin artması, ve yazılı ve sözlü metinlerin derlenmesinde yaşanan hızlı büyüme sayesinde derlem dilbilimde müthiş bir canlılık yaşanmaktadır. Gries’in de (2009) belirttiği gibi derlem dilbilim, dilbilimin en hızlı büyüyen yöntemlerinden biridir ve aynı zamanda gittikçe de daha fazla dili kapsayarak ‘çokdilli’ olmaktadır (McEnery ve Wilson, 2001). McEnery ve Wilson (2001) derlem dilbilimi ‘gerçek hayatta’ kullanılan dilden alınan örneklere dayanılarak yapılan dil çalışması olarak tanımlamaktadır. Derlem dilbilim söylev çalışmaları, gramer, dil eğitimi, kültür, sosyodilbilim, anlambilim ve sözcükbilim çalışmaları gibi pek çok alanda kullanılmaktadır. Sözcükbilim çalışmaları dil ile ilgili daha güncel veriler sağlayarak sözlüklerin eskisine kıyasla çok daha hızlı bir şekilde hazırlanmasını ve yenilenmesini sağlamaktadır. Dilbilimciler uzun zamandan beri bir arada olma eğiliminde olan ardışık kelimeleri inceliyorlar (Strunkyte & Jurkunaite, 2008). Bu çalışmada sözcük örüntüsü denilen ve sürekli tekrar eden belirli bir kelime grubu incelenmektedir. Sözcük örüntüsü, ‘küme’ ya da ‘yığın’ da denilen birkaç kelimenin yan yana gelmesi ile oluşan kelime gruplarıdır ve dili akıcı bir şekilde kullanmanın önemli bir bölümü bu yapılara hakim olmaktan geçer (Hyland, 2008a). Bu çalışma anadili İngilizce olan ve İngilizce’yi yabancı dil olarak öğrenen Türk ve Japon öğrenciler tarafından tartışma türünde yazılmış kompozisyonlarda en çok kullanılan ‘tutum sözcük örüntüleri’ni (TSÖ) bulmayı amaçlamıştır. Aynı zamanda, bu kelime grubunun yapısal ve işlevsel olarak hangi özellikler gösterdiği ve bu özelliklerin anadili

İngilizce olan ve İngilizce'yi yabancı dil olarak öğrenen öğrenciler arasında nasıl bir farklılık gösterdiğini bulmak hedeflenmiştir. Çıkan sonuçlara göre anadili İngilizce olan katılımcılar TSÖ'yü en az kullanırken en fazla kullanan grup Japon öğrenciler olmuştur. TSÖ'lerin yapısal ve işlevsel kullanımları her grupta farklılık göstermiştir.

Anahtar kelimeler: Derlem dilbilim, tutum sözcük örüntüleri, karşılaştırmalı aradil çalışması

ABSTRACT

STANCE LEXICAL BUNDLES IN ACADEMIC L2 WRITING: A CONTRASTIVE INTERLANGUAGE ANALYSIS

Meltem MUŞLU

PhD Dissertation, English Language Teaching Department
Supervisor :Assoc. Prof. Dr. Cem CAN
December, 2014, 92 Pages

Corpus linguistics is ‘’a linguistics methodology which is founded on the use of electronic collections of naturally occurring texts, viz. corpora’’ (Granger, 2002, p.4). In the last five decades, with the acceleration in the development of computer technology in terms of processing speed and storage capacity of electronic data, and the structured compilation of written or transcribed speech texts, corpus linguistics (CL) has undergone a remarkable revival. As Gries (2009) states, CL is one of the fastest-growing methodologies in linguistics, which is also a lively and increasingly multilingual subject (McEnery& Wilson, 2001). McEnery and Wilson (2001) define CL as the study of language based on examples of ‘real life’ language use. Corpus linguistics is used in many areas, such as speech research, grammar, applied linguistics, culture, sociolinguistics, semantics and lexical studies. Lexical studies help producing and revising dictionaries much more quickly than before by providing more up-to-date information about the language. Linguists have been interested in the sequences of words that tend to occur for a long time (Strunkyte&Jurkunaite, 2008). The present study examines a particular type of recurrent chunks: lexical bundles. Lexical bundles are multi-word expressions referred to as “clusters” or “chunks” and an important component of the fluent linguistic production depends on the control of them (Hyland, 2008a). This study aimed at finding out the most common stance lexical bundles used in argumentative essays written by native English speakers and Turkish and Japanese EFL learners. It also aimed to find the structural and functional characteristics of these bundles, and to find out how similar these structures are used by the native speakers and EFL learners. The results showed that native speakers use lexical bundles least.

Japanese EFL use these chunks most frequently. The functions and the structures of lexical bundles vary in each group.

Keywords: Corpus linguistics, stance lexical bundles, contrastive interlanguage analysis

ACKNOWLEDEMENTS

There are many people who I am deeply indebted for the help and support they have given me for the completion of this thesis.

First of all, I would like to thank and express my deepest gratitude to my thesis advisor, Assoc. Prof. Dr. Cem CAN for his invaluable guidance, support, and encouragement during the completion of this thesis. His energy, optimism and positive comments shed a light to my way whenever I felt down.

My deep thanks also go to Prof. Dr. Hatice SOFU. This thesis would never be completed without her patience and guidance. She sacrificed her time and provided me feedback whenever I needed help; even on a train to Warsaw. I also would like thank Assoc. Prof. Dr. Ahmet DOĞANAY. He not only taught statistics but also how to be a good academician with his unlimited patience and kindness. My thanks also go to Assoc. Prof. Dr. Türkay BULUT and Assist. Prof. Dr. Jülide İNÖZÜ for accepting to be my committee members and for their insightful comments at the end of my study.

I was very lucky to have my fiends Celeste SHOPE and Laurel ANN with me when writing my thesis. It was very mind-opening to talk to Celeste during the analysis of my data. She spent hours to analyze my data with me. I am very thankful to Laurel for editing my thesis and giving feedback.

I am also very grateful to my family and all my friends for supporting me all the time. My family- without their presence, support, encouragement, understanding and love, this thesis would never be completed. Seda BANLI- thank you for adding joy to my exhausting days in Adana.

TABLE OF CONTENTS

Page

ÖZET iii
ABSTRACT v
ACKNOWLEDGEMENTS vii
LIST OF ABBREVIATIONS viii
LIST OF TABLES ix
LIST OF FIGURES x
LIST OF CHARTS xi
LIST OF APPENDICES xii

CHAPTER I
INTRODUCTION

1.0. Introduction 1
1.1. Background of the Study 1
1.2. Statement of the Problem 2
1.3. Aim of the Study 2
1.4. Research Questions 2
1.5. Operational Definitions and Key Terms 3
1.6. Limitations of the Study 4

CHAPTER II
REVIEW OF LITERATURE

2.0. Introduction 5
2.1. Corpora and Corpus-based Studies 5
2.1.1. Corpus in Applied Linguistics 6
2.1.2. Learner Corpora 8
2.2. Lexical Bundles 9
2.2.1. Structural Types of Lexical Bundles 10
2.2.2. Functional Categories of Lexical Bundles 10
2.2.2.1. Stance Expressions 11
2.2.2.2. Discourse Organizers 12

2.2.2.3. Referential Expressions 13
2.3. Significance of Lexical Bundles 13
2.4. Major Studies on Lexical Bundles 14

CHAPTER III

METHODOLOGY

3.0. Introduction 19
3.1. Data Collection Tools and Participants 19
3.2. Data Analysis 23
3.2.1. Cut-off Point 25
3.2.2. Word Smith 26
3.2.3. Type-token Ratio 28
3.2.4. Log Likelihood 29
3.2.5. Analyzing Functions and Structures of Stance Lexical Bundles 30

CHAPTER IV

RESULTS AND DISCUSSIONS

4.0. Introduction 33
4.1. Overall Results of the Study 33
4.1.1. Log Likelihood Analysis Results 34
4.1.2. Type-token Ratio Results 36
4.1.3. Forty Most Common Lexical Bundles 38
4.2. Functions of Stance Lexical Bundles 41
4.3. Structures of Stance Lexical Bundles 46
4.4. Comparison of the Three Corpora 49

CHAPTER V

CONCLUSIONS

5.0. Introduction 52
5.1. Summary of the Study 52
5.2. Implications and Recommendations for Language Learning and Teaching 56
5.3. Limitations of the Study 58

5.4. Future Research 58

REFERENCES.....60
APPENDICES.....68
CURRICULUM VITAE.....90

LIST OF ABBREVIATIONS

EFL : English as a Foreign Language

ICLE : International Corpus of Learner English

JPICLE : Japanese International Corpus of Learner English

LBs : Lexical Bundles

LL : Log Likelihood

LOCNESS : Louvain Corpus of Native English Essays

L1 : First Language

L2 : Second Language

SLBs : Stance Lexical Bundles

TICLE : Turkish International Corpus of Learner English

T/t :Type Token Ratio

LIST OF TABLES

Page

Table 1. Top Ten Essay Topics in ICLE 20
Table 2. Turkish Sub-corpus 21
Table 3. Japanese Sub-corpus 22
Table 4. LOCNESS Selected Components 23
Table 5. Research Type and Stages of the Study 24
Table 6. Sample Table for Type-token Ratio 29
Table 7. Frequencies and Percentages of Stance Lexical Bundles 33
Table 8. LL Ratio of TICLE and LOCNESS 35
Table 9. LL Ratio of JPICLE and LOCNESS 35
Table 10. LL Ratio of JPICLE and TICLE 36
Table 11. Overall Comparative Frequency Distribution of Epistemic Bundles in General In Three Corpora 36
Table 12. Overall Comparative Frequency Distribution of Epistemic and Attitudinal/Modality Stance Lexical Bundles in Three Corpora 37
Table 13. Most Common 40 Stance Lexical Bundles 38
Table 14. Functions of Stance Lexical Bundles 42
Table 15. Use of Self Mention in Corpora 44
Table 16. Structures of Stance Lexical Bundles in Detail 47
Table 17. Structures of 40 Most Common LBs in Detail 48

LIST OF FIGURES

Page

Figure 1. Categories of Lexical Bundles.....9
Figure 2. Structural Types of Lexical Bundles 10
Figure 3. Functional Categories of Lexical Bindles. 11
Figure 4. A Screenshot of WordSmith Main Page 26
Figure 5. A Screenshot of WordList Page in WordSmith Tools 27
Figure 6. A Screenshot of Concord Tool in WordSmith Tools 28
Figure 7. Contingency Table.....29
Figure 8. Categories of Epistemic Stance Bundles.....31
Figure 9. Functional Categories of Stance Lexical Bundles Used in This Dissertation 32

LIST OF CHARTS

Page

Chart 1. Functions of Lexical Bundles in general 41

Chart 2. Structures of Stance Lexical Bundles in general 46

LIST OF APPENDICES

Page

Appendix A. LOCNESS Results 68
Appendix B. TICLE Results 73
Appendix C. JPICLE Results 81

CHAPTER I

INTRODUCTION

1.0. Introduction

Corpus linguistics is ‘a linguistics methodology which is founded on the use of electronic collections of naturally occurring texts, viz. corpora’ (Granger, 2002, p.4). In the last five decades, with the acceleration in the development of computer technology in terms of processing speed and storage capacity of electronic data, and the structured compilation of written or transcribed speech texts, corpus linguistics (CL) has undergone a remarkable revival. As Gries (2009) states, CL is one of the fastest-growing methodologies in linguistics, which is also a lively and increasingly multilingual subject (McEnery& Wilson, 2001). McEnery and Wilson (2001) define CL as the study of language based on examples of ‘real life’ language use. A corpus consists of a data bank of natural texts which are compiled from writing or a transcription of recorded speech (Krieger, 2003).

1.1. Background of the Study

Corpus linguistics is used in many areas, such as speech research, grammar, applied linguistics, culture, sociolinguistics, semantics and lexical studies. Lexical studies help producing and revising dictionaries much more quickly than before by providing more up-to-date information about the language. It also has an important role in studying the frequencies of different morphological variants and the productivity of different morphemes (McEnery& Wilson, 2001).

Linguists have been interested in sequences of words that tend to occur for a long time (Strunkyte&Jurkunaite, 2008). These groups of words that frequently occur together in a language have been studied and described under different labels, such as clusters, chunks and lexical bundles. The present study examines a particular type of recurrent chunks: lexical bundles. Lexical bundles are multi-word expressions referred to as “clusters” or “chunks” and are an important component of the fluent linguistic production which depends on the control of them (Hyland, 2008a).

1.2. Statement of the problem

Although formulaic language has been the focus of linguistic inquiry for several decades, more compositional subclasses of formulaic sequences differing from idioms and set expressions in their structural and functional characteristics are quite ignored (Nekrasova, 2009). Studies on formulaic language largely focus on the academic written and spoken genres of the English language (Cortes 2002, Hyland 2008a, Hyland 2008b, Kim 2009, Neely and Cortes 2009, Bal 2010), focus on teaching and textbook language (Biber, Conrad & Cortes 2004), and focus on the comparison of the use of lexical bundles by native and non-native speakers of English (Cortes 2004, Juknevičienė 2011). However, none of these studies compare the use of lexical bundles by learners from different native language backgrounds. As Hyland (2008 b) points out, more work with different disciplines, genres, and first language groups should be done, since as Granger (2002) points out, an improved understanding of learner output illuminates all aspects of pedagogy from tasks to curriculum. Therefore, this study aims at comparing the use of lexical bundles by native and non-native speakers of English.

1.3. Aim of the study

The purpose of the study is twofold. First, it aims at exploring the forms, structures and functions of stance lexical bundles used by native speakers of English, andby Turkish and Japanese learners studying English as a Foreign Language (EFL). The second aim is to find out whether there is a difference in the use of lexical bundles between the native speakers of English and the Turkish and Japanese EFL learners.

1.4. Research questions

This dissertation sought answers for the following questions:

1. What are the most common stance lexical bundles found in argumentative essays written by
    a. native English speakers?
    b. Turkish and Japanese EFL learners?
2. What are the structural and functional features of the stance lexical bundles used by native English speakers, and Turkish and Japanese EFL learners?

3. To what extent are the stance lexical bundles used by the Turkish and Japanese EFL learners similarto the ones used by the native speakers?
4. Do the differences/similarities regarding the use of lexical bundles stand from the learners' first language or is this an interlanguage property?

1.5. Operational definitions and key terms

Computer Learner Corpus (CLC): Electronic collection of authentic texts produced by foreign or second language learners. (McEnery and Wilson, 2001, p.177)

Concordance: 'A comprehensive listing of a given item in a corpus (most often a word or phrase), also showing its immediate context'' (McEnery& Wilson, 2001, p.197)

Contrastive Interlanguage Analysis (CIA): A method involving comparisons between native speakers and learners (L1 vs. L2), and between different learner groups (L2 vs. L2) (Granger, 1996).

Corpus: ''A corpus is a collection of texts (a 'body' of language) stored in an electronic database. Corpora are large bodies of machine readable texts containing thousands or millions of words'' (Baker, Hardie&McEnery, 2006, p. 48).

Corpus Linguistics (CL): ''A linguistics methodology which is founded on the use of electronic collections of naturally occurring texts, viz. corpora'' (Granger, 2002, p.4).

International Corpus of Learner English (ICLE):A learner corpus containing of argumentative essays written in English by learners from 16 different mother tongue backgrounds (Granger, Dagneaux&Meunier, 2009).

Japanese International Corpus of Learner English (JPICLE):A learner corpus containing the Japanese EFL learners' written argumentative essays (Granger et al, 2009).

Learner Corpus (LC):A corpus containing written or spoken texts of second language (L2) learners.

Lexical Bundles (LB): Lexical bundles are the most frequently occurring sequences of words which are usually not idiomatic in meaning—the meanings are transparent from the individual words— and they are usually not complete grammatical structures (Biber, 2006).

Louvain Corpus of Native English Essays (LOCNESS):A native reference corpus containing British and American Students' written essays (Granger et al, 2009).

Stance Lexical Bundles (SLB): Stance bundles express attitude or assessment of certainty. They provide a frame in which one can interpret the coming information (Biber, 2006).

1.6. Limitations of the study

The present study is limited to the size of the three corpora, therefore the results of the study are limited to the analysis of them; LOCNESS as native English corpus, TICLE as a learner corpus of Turkish EFL learners, and JPICLE as a corpus of Japanese EFL learners. In addition, the study has been carried out by limiting its scope to overuse and underuse of lexical bundles by EFL learners; misuse of the structures are not included in the analysis.

CHAPTER II

REVIEW OF LITERATURE

2.0. Introduction

This chapter will provide background information for the current study by first giving information on the meaning of corpora and how corpus-based studies are conducted. Then, the description and characteristics of lexical bundles will be explained in detail.Significance of lexical bundles, literature review on these expressions and recent corpus-based studies related to them will follow these.

2.1. Corpora and Corpus-based Studies

Corpus is a Latin word meaning ‘body’. Although it means a body or collection of texts in any form, today linguists usually mean a collection of computer-readable texts compiled using a clearly delineated set of design criteria (Cheng,Warnen&Xun-Fen, 2003). McEnery and Wilson (2001) define corpus linguistics (CL) as the study of language based on examples of ‘real life’ language use. A corpus consists of a data bank of natural texts which are compiled from writing or a transcription of recorded speech (Krieger, 2003).

As McEnery& Wilson (2001) state, CL is not a branch of linguistics in the same sense as syntax, semantics, and so on, which concentrate on describing or explaining some aspect of language use. CL is not an aspect but a methodology; therefore, a corpus-based approach can be taken into many aspects of linguistic inquiry. Syntax, semantics and pragmatics are only three of the areas of linguistic inquiry which use a corpus based approach. There are four characteristics of corpus based analysis. Firstly, it is empirical; actual patterns of use in natural texts are analyzed. Secondly, it utilizes a large and principled collection of natural texts. Thirdly, extensive use of computers for analysis is essential, and lastly, both quantitative and qualitative analysis is used (Biber et al, 1999).

A corpus is designed for a particular purpose and the types of corpus depend on this particular purpose. As Hunston (2002) states, the most commonly used corpus types are specialized, general, comparable, parallel, pedagogic, historical, monitor and learner corpus-which is the main instrument of this dissertation. Learner corpus is a collection

of student texts. Its aim is to identify in what respects learners differ from each other and from the native speakers.

Corpus linguistics is used in many areas, such as speech research, grammar, applied linguistics, culture, sociolinguistics, semantics and lexical studies. Lexical studies help produce and revise dictionaries much more quickly than before and provide more up-to-date information about the language. They also have an important role in studying the frequencies of different morphological variants and the productivity of different morphemes (McEnery& Wilson, 2001).

2.1.1. Corpus in Applied Linguistics

Corpus studies have a crucial role in applied linguistics. As McEnery& Wilson (2001) state, corpus examples are important in language learning since they expose students to the kinds of sentences and vocabulary that they will encounter in reading genuine texts in the language or in using language in real communicative situations. They also have an important role inlooking at the existing language teaching materials critically.

One application of corpus studies in language teaching is teaching languages for specific purposes. This refers to the domain-specific teaching of a language for a particular area of application, such as the teaching of medical English to medical students. Other applications of corpus in applied linguistics are computer-assisted language learning and translation teaching. Using corpora in language teaching has a positive effect on students' performance as found in the study by McEnery, Baker & Wilson (1995; cited in McEnery&Wilson 2001). The availability of multilingual parallel corpora makes translation teaching easier by providing side-to-side examples of style and idiom in more than one language and of being able to generate exercises in which students can compare their own translations with an existing professional translation or original (McEnery& Wilson, 2001).

It also has applications in classroom regardless of the students' language levels. Corpora can be used as a tool to correct students' mistakes. After finding out the students' mistakes, providing the appropriate type of corpora and teaching methods, the mistakes may be corrected by the teachers or the students themselves. Some current studies applying corpus in classroom teaching to see its effectiveness are summarized below.

Rodriguez and Moreno (2009) conducted a study to provide a strategy to teach idioms through a corpus based on Disney movies. Their aim was to provide students with the necessary tools to make sense of incongruous expressions. Their aim was also building up the students' metaphoric competence while making learning more enjoyable. In their study, they give examples and explain how they used a corpus based on Disney movies to teach idioms.

Liu and Jiang (2009)conducted a detailed study by collecting a variety data; students' corpus search projects and reflection papers, teachers' lesson plans and teaching journals, and a post study assessment survey. They found that use of corpora and lexicogrammar make learning more effective by enhancing students' language awareness, improving their command of lexicogrammatical rules and usage patterns, increasing their appreciation of context in language use and promoting discovery learning.

Gajic (2010) conducted a study to show how business related periodical articles can be used in teaching. The participants in the study were advanced level EFL students. In the study two corpora, one with articles in English and one with articles in Serbian, were used to analyze the languages linguistically. Afterwards, teaching materials based on these findings were prepared.Gajic concludes that the study provides valuable information beneficial for language pedagogy and course design.

Girgin (2011) conducted a study to see the effectiveness of using corpus-based activities on lower level EFL students' learning of English grammar. His purpose was to determine whether lower level EFL learners would be able to use corpus-based activities effectively to learn the target structures he determined and to find out the students' attitudes towards these activities. He found that the students were able to use the corpus-based activities effectively. He also found that although the students held somewhat neutral attitudes towards using corpus-based activities in the questionnaire, the student interviews revealed more positive attitudes towards these activities.

Sahillioğlu, Şahinkayası and Şahinkayası (2011) conducted a study to compare two corpus-driven teaching approaches for teaching/learning grammar. Their aim was to compare the effectiveness of a small and a large corpus. They found that using small and large corpora did not make a significant difference on the students' progress in learning grammar structures; however, they also found that the group taught with a small corpus was more successful in forming grammatical structures.

Jafarpour, Hasmeidan and Alipour (2013) compared the effects of the corpus-based approach (concordance-based) with the effects of the traditional approach on learning collocations of near-synonymous pairs. They found that concordance-based methods had more effect on collocation knowledge of the learners. The participants in the concordance-based method had better comprehension and production of the collocations taught.

As seen above, corpus-based teaching is one of the applications of using corpora in classroom teaching. Another application of corpora in classroom teaching is learner corpora, which is one of the data collection tools used in this dissertation.

2.1.2. Learner Corpora

Computer learner corpora is defined as ‘electronic collections of spoken or written texts produced by foreign or second language learners in a variety of language settings’ (Granger et al, 2002; p. vii). The main purpose of compiling a learner corpus is to gather natural interlanguage data to describe and analyze learner language (Granger, 1998).As Granger (2002) states, learner corpus research has created an important link between the two previously disparate fields of corpus linguistics and foreign/second language research. Learner corpora provide a new type of data which can inform thinking both in second language acquisition research, which tries to understand the mechanisms of foreign/second language acquisition, and in foreign language teaching research, whose aim is to improve the learning and teaching of foreign/second languages. Learner corpus research opens up exciting pedagogical perspectivesin different areas of pedagogy:materials design, syllabus design, language testing, and classroom methodology(Granger, 2003).With the help of the linguistic software tools, from simple ones which search, count and display to the most advanced ones which provide sophisticated analyses of the data, the data can be analyzed after being computerized. Until quite recently, data collection and analysis required tremendous time and effort on the part of the researcher; therefore, the samples of learner data used in SLA studies have traditionally been rather small, but technological progress has made it perfectly possible to collect learner data in large quantities, store it on the computer and analyze it (semi)automatically using currently available linguistic software (Granger, 2002). As Mukherjee (2006, 20) states ‘Learner corpus research is an area which is immediately relevant to the teaching and learning of languages, but its full

potential has not yet been exploited in language pedagogy.' Although there have been different studies done using learner corpus, more work should still be done. Some of the studies using corpora focus on different words classes (in different languages), such as verbs or adverbs, and compare the variation in written and spoken registers (Newman & Rice, 2004; de Klerk, 2005; Louwerse et al, 2004; Anğ, 2006; Başal, 2006; Guo, 2006 Şanal, 2007; Babanoğlu, 2012). Yet, some studies focus on learner corpus and classroom teaching, and preparing corpora in different languages, such as Japanese, Chinese, and Arabic (Shih, 2000; Tono et al, 2001; Reder& Harris, 2003; Tagnin, 2006; Abuhekema et al, 2008; Belz 2008; Hua, 2008). Although many studies are done on corpora, not many studies have focused on the use of lexical bundles.

2.2. Lexical Bundles

Lexical bundles, which will be referred as LBs hereafter, are the most frequently occurring sequences of words which are usually not idiomatic in meaning;the meanings are transparent from the individual words, and are usually not complete grammatical structures (Biber, 2006).They are combinations of three or more words which are identified empirically in a corpus of natural language; for instance,*take a look at, I don't know, on the other hand*, and*as a result of* (Cortes, 2006). They are extended collocations which appear more frequently than expected by chance, helping to shape meanings in specific contexts and contributing to our sense of coherence in a text (Hyland, 2008 b).

Lexical bundles are classified into two categories by Biber et al. (2004); structurally and functionally.

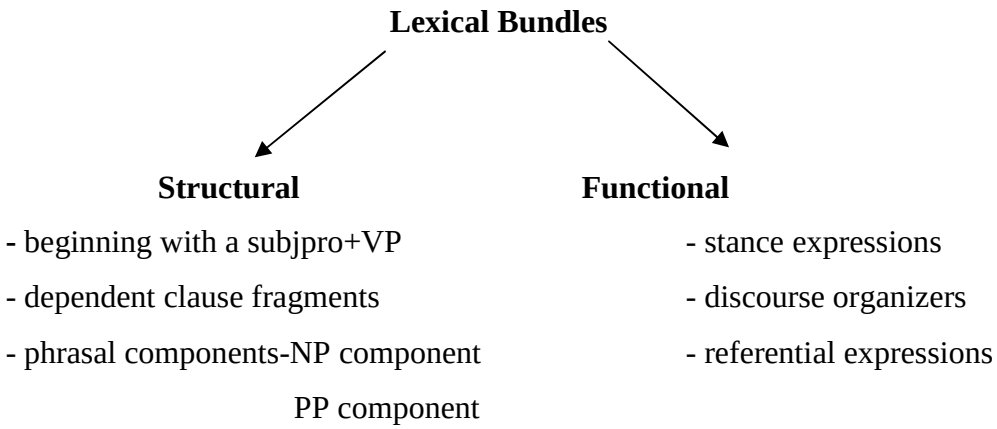


Figure 1. Categories of lexical bundles

2.2.1. Structural Types of Lexical Bundles

Although lexical bundles are not usually complete structural units, they have strong grammatical correlates. There are three major structural types of lexical bundles. These types are presented in the figure below.

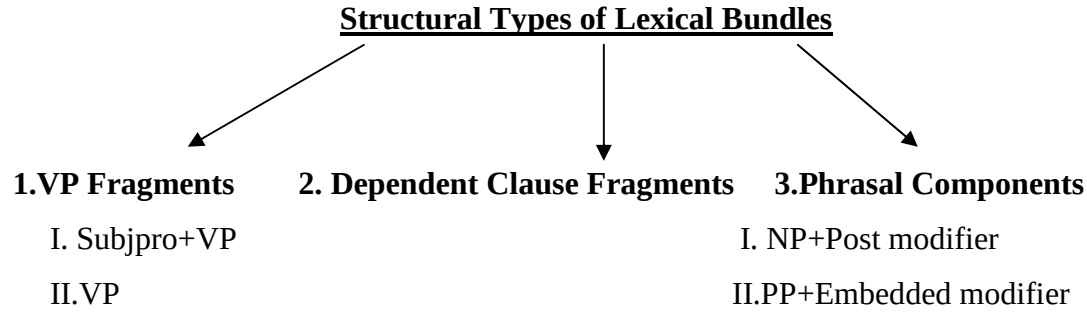


Figure 2. Structural types of lexical bundles

The first major structural type incorporates verb phrase fragments. These bundles often begin with a subject pronoun followed by a verb phrase, such as *I'm notgoing to, it's going to be, that's one of the,* and *this is a*. These bundles can also begin directly with a verb phrase, for instance *is going to be, take a look at* or question fragments, such as*are you going to* and *how many of you*.

The second major structural type is similar to the first category in that it incorporates verb phrase elements, but these bundles also incorporate dependent clause fragments, for example *I want you to, if we look at,* and*what I want to*.

The last major structural type of lexical bundles includes only phrasal components. Many of these bundles consist of noun phrase components, usually ending with the start of a post modifier, such as *the end of the, a little bit about,those of you who, the way in which*. Other bundles of this type consist of prepositional phrase components with embedded modifiers, for instance *of the things that* and *atthe end of*.

2.2.2. Functional Categories of Lexical Bundles

The functional categories of lexical bundles include stance expressions, discourse organizers, and referential expressions with several sub-categories under each of these groups. These bundles are presented in Figure 3 below.

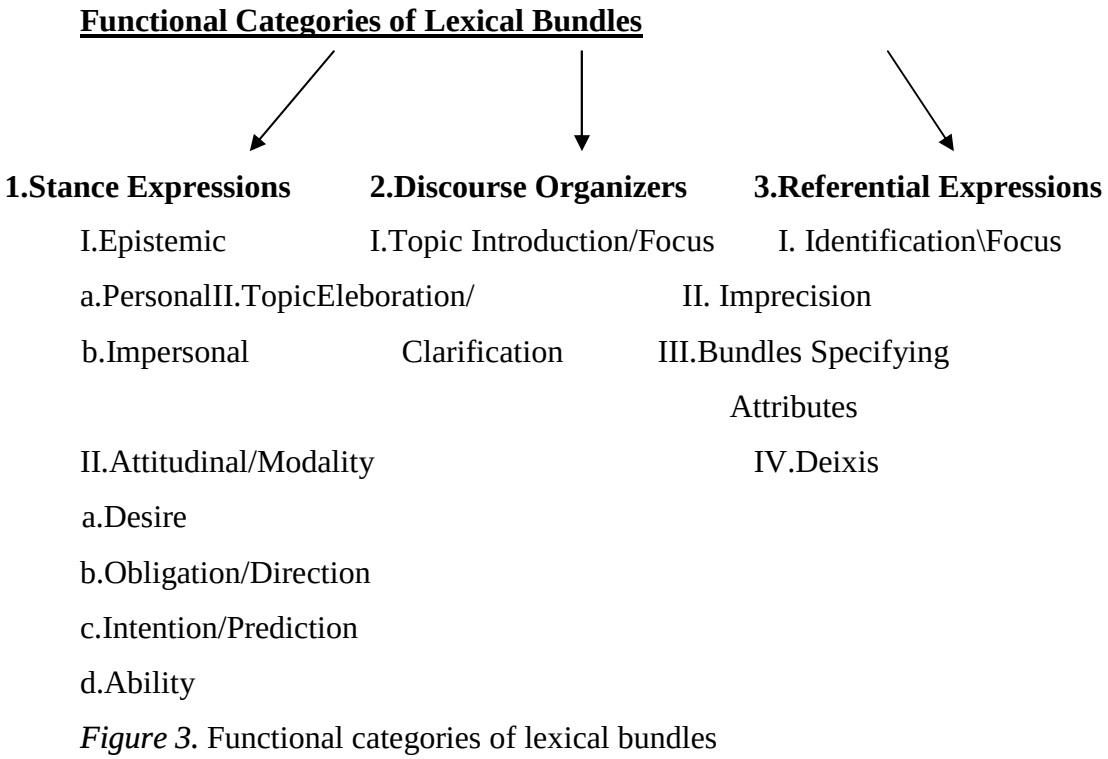


Figure 3. Functional categories of lexical bundles

2.2.2.1. Stance Expressions

Stance bundles express attitude or assessment of certainty. They provide a frame in which one can interpret the coming information, for instance *I don't know what* and *if you want to.* Stance bundles can be personal and impersonal. Personal stance bundles overtly refer to the speaker/writer with the subject 'I'; for instance *I don't know what.* Impersonal stance bundles are similar to personal stance bundles, but they do not attribute the speaker/writer overtly; for example, *it is possible to, can be used to*. Stance bundles convey two major kinds of meaning; epistemic and attitude/modality.

Epistemic stance bundles comment on the knowledge status of the information (certain, uncertain, probable/possible), such as *I don't know if,* and *I don't think so.*

There are two sub-categories of epistemic stance bundles: Personal and impersonal.Personal epistemic bundles can express certainty and uncertainty but most of them express only uncertainty. *I don't know what, I don't think so,* and *I think it was* are examples of personal epistemic bundles.Impersonal epistemic bundles express certainty. *Are more likely to,* and *the fact that the* are in this group.

Attitudinal/modality stance bundles express speakers' attitudes towards the actions, for instance *I want you to* and *I am not going to*. There are four subcategories of this group: desire, obligation/directive, intention/prediction, and ability.

Only personal expressions of stance are included in the desire bundles. They frame self-motivated desires and wishes or ask about another participant's desires. *I would like to, what I want to do* are some of the bundles in this group.

Obligation/directive bundles are also personal stance bundles, but they differ from them in that they have a second person pronoun; 'you' instead of 'I'. Some examples in this group are *I want you to, you have to do,*and *you need to know.* A few of them are impersonal with no pronoun, such as *it is important to.*

Intention/prediction bundles are overtly personal. They express the speaker/writer's future intentions openly. *Is going to be, what we are going to* are some of the bundles in this category.

Some stance bundles express ability, such as *to be able to* and *to come up with.* These bundles are mostly used in the classroom discourse

2.2.2.2. Discourse Organizers

Discourse organizers express relationships between prior and coming discourse. They allow for introducing new topics and elaborate on given topics,such as *I want to talk about, if you look at.* They have two major functions: topic introduction/focus and topic elaboration/clarification.

Topic introduction/focus bundles signal that a new topic will be introduced. They also express intention and desire, but their special function is to announce the new topic. *I want to talk about, what I want to do, if we look at* are the bundles in this group.

Topic elaboration/clarification bundles are used when additional explanation or clarification are needed. For instance, *you know, I mean.* Some of them are used for explicit comparison and contrast: *as well as the, on the other hand.*

Identification/focus bundles are also considered as discourse bundles, although Biber, Conrad and Biber (2004) classify them as referential expressions. Mostly, they have a discourse organizing function; for instance, *that's one of the* is used after a lengthy explanation to summarize or emphasize a point. Similarly, *those of you who* identifies a subgroup. They can also be used to introduce an idea, such as *one of the things.*

2.2.2.3. Referential Expressions

Referential bundles make direct reference to the textual context or the physical or abstract entities. They specify an attribute of something as being important, for example *that's one of the, the nature of the.* There are four sub-categories of referential expressions: identification/focus, imprecision, specification of attributes, and time/place/text reference.

Imprecision bundles indicate imprecise reference either to indicate a specified reference is not exact or to indicate that there are additional references of the same type; such as *or something like that,*and *things like that.*

The second subcategory of referential bundles identifies specific attributes of the head noun they precede. *Have a lot of, a little bit of,a little bit about* can be considered in this group. Some of these bundles describe the size and form of the head noun they precede: *the size of the, in the form of.* Some of them describe abstract characteristics,such as *the nature of the.*

Time/place/text deixis bundles refer to particular time, place and location in the text. Text deixis bundles are only common in written discourse, such as *as shown in Figure.* Other deixis bundles can be *in the United States, at the end of the.*

2.3. Significance of Lexical Bundles

Gaining sensitivity to expert users' preferences for certain sequences of words is required in gaining the control of a new language (Hyland, 2008a). Identifying fixed phrases provides benefits for both language teachers and learners since fluent linguistic production depends on the control of the lexical bundles. Identifying them may help learners acquire the specific rhetoric practices of the texts they are asked to write (Hyland, 2008b). As Hyland (2008a, 60) states 'the study of clusters offers insights into a crucial, and often overlooked dimension of language use, providing a better understanding of the ways writers employ the resources of English in different contexts, and with the potential to inform advanced academic literacy instruction'.

The ability to express epistemic stance appropriately in English is considered as crucial for successful interaction between the writer and the reader. Coates (1995, p. 55) argues that epistemic modality is "concerned with the speaker's assumptions, or assessment of possibilities, and, in most cases, it indicates the speaker's confidence or lack of confidence in the truth of the proposition expressed".Epistemic stance can also

be used by writers to realize diverse social and pragmatic functions (Holmes, 1982; Coates, 1990).

2.4. Major Studies on Lexical Bundles

Lexical bundles were first defined in the Longman Grammar of Spoken and Written English (Biber, 2006) and after that, an increasing number of studies have been done on them. Considering their significance in language learning and teaching, researchers have been investigating the use of lexical bundles in classroom teaching and textbooks, and in professional writers' and students' writings.

First studies on LBs focused on the identification of these units and a description of their patterns of occurrence in different written and spoken registers in English.The core study on LB was done by Biber (2006). The study was based on Longman Spoken and Written English Corpus (LSWE), which is a large corpus of both American and British English conversation and academic prose. The results of the study showed that a large number of different lexical bundles are used both in conversation and academic prose.

Another major study on lexical bundles is also done by Biber and his colleagues (2004). In this study they looked at the use of LBs in university classroom teaching and textbooks by using LSWE corpus again. It is found that university lectures use twice as many LBs than conversation and four times as many LBs as textbooks. Another significant finding of the study was the linguistic features they found in their corpora.

Conrad &Biber (2004) also analyzed the LBs in academic prose and conversation. The study revealed that different registers rely on different sets of lexical bundles, and that the bundleshave important discourse functions that fit the context and purposes of the registers inwhich they are common. Their use is not accidental; the frequency and functions of LBs showed that speakers and writers use them regularly in building discourse. They concluded that further study is needed particularly from a psycholinguistic perspective and in more registers.

The use of LBs in non-academic and core instructional registers is also looked at (Biber&Barbieri, 2007). It was found that LBs were also very common in instructional written registers, such as course syllabi. This finding is in contrast with previous studies showing that LBs were more common in speech than in writing.

A study by Nesi and Başẗurkmen (2006) also focused on the use of LBs in academic lectures. This study shows that lexical bundles can play a discourse signaling role in lectures; therefore, raising students' awareness on the use of LBs is important.

In her study, Cortes (2002) analyzed the use of LBs in freshman composition. She found that expression serving as locative or temporal markers created redundancy in students' compositions. She concluded that students' writings at different levels and different disciplines should be evaluated both structurally and functionally.

Hyland (2008a)investigated the forms, structures and functions of LBs in a 3.5 million word corpus of research articles, doctoral dissertations and master's theses in electrical engineering, biology, business studies and applied linguistics by identifying the most frequent 4-word bundles. The results of the study revealed that LBs are central to the creation of academic discourse and they offer an important means differentiating written texts by discipline.

Another primary study done in 2008 (b) by Hyland again used the research articles, doctoral dissertations and master's theses. The results of the study showed that professional academics, doctoral students and master's level students rely on different resources to develop their arguments, to be more persuasive and to be credible. Also, these three uses are characterized by different clusters.

Strunkyte&Jurkunaite (2008) compared the use of LBs in the articles published in the humanities and natural sciences. They found that research articles in humanities contain a large stock of lexical bundles than natural sciences. Also, in humanities, stance and referential bundles occurred more frequently; whereas in natural sciences, discourse organizing and referential bundles occurred more frequently.

Herbel-Eisenmann, Wagner & Cortes (2010)also looked at LBs, but they narrowed their research to stance lexical bundles. They compared a corpus from a secondary mathematics classroom with two other corpora; one from university classrooms and one from conversations. They concluded that teacher stance is significant to teachers' and children's' experience in mathematics classroom. Also, there are particular characteristics of positioning in math classes that are fundamental to students' experiences of mathematics.

Jablonkai (2010) compiled the English EU Discourse Corpus to find out the frequent lexical items used in European Union context. The results of the study showed thata fairly large proportion of EU texts are made up of formulaic patterns. The findings

provide a more clear understanding of the special characteristics of EU discourse and also they can serve as a basis for course and materials design for EU English courses.

Hernandez (2013) looked at the forms, structures and functions of 4-word bundles in three corpora of spoken English. One of the corpora consisted of the data gathered from native speakers of English and the remaining two were the corpora of non-native speakers of English. The study focuses on the overall distribution of bundles, their typical structures and functions. The findings of the study showed significant differences in the types of lexical bundles used by native and non-native students, as well as in their structure and function.

In addition to these studies focusingon the identification of LBs and a description of their patterns of occurrence in different written and spoken registers in English, a few studies focused on a more pedagogical aspect of them.

In her first study, Cortes (2006) focused on the teaching of LBs to a group of university students in a writing-intensive history class. Pre- and post-instruction analyses were conducted on students' class assignments in order to identify the use of lexical bundles. The comparison of pre-and post-instruction showed no difference on the production of lexical bundles, but an increase in students' awareness of and interest in these expressions was seen.

Neely & Cortes (2009) focused on the use of LBs in order to compare the use of bundles by instructors to that of students. The findings of the comparison were used as the basis for the design of a series of academic listening lesson plans which focus on those bundles that most often occur in academic lectures and the functions they perform in that context. After a series of classroom teaching, they concluded that in order to design lessons and materials for the ESL/EFL classroom, corpus-based research findings in conjunction with publicly available corpora and concordance programs can be used. Bundles should be taught by presenting the complete spectrum of their functions in context by analyzing these functions in discourse similar to the one they encounter daily in their academic lives.

Comparison of different language backgrounds has a crucial role in applied linguistics since it would give insights in what to teach and how to teach them. Below are some studies done to determine and compare LBs found in different languages.

In her study, Cortes (2008) compared the LBs in the published history articles written in English and Spanish. It was found that the number of lexical items was

different, but there was a certain degree of agreement in the expressions used in the two corpora.

Nekrasova (2009)also compared the use of lexical bundles by first and second language speakers of English. The study examined whether L1 and L2 English speakers display any knowledge of lexical bundles as holistic units and whether their knowledge is affected by the discourse function of the LBs.It was found that the participants' knowledge differed for specific lexical bundles and they knew more discourse-organizing bundles than referential bundles.

Bal (2010)investigated the use of lexical bundles in research articles written in English byTurkish scholars. She identified and analyzed the LBsboth structurally and functionally. The study showed that the LBs found have structural correlates as wellas strong functional features that help to construct discourse in academic writing.

Tenuta et al. (2012) compared the use of LBs in two corpora, Corpus of Brazilian Learners of English (CABrI) and Louvain Corpus of Native English Essay (LOCNESS). The analysis showed that epistemic modals were used extensively in both corpora, with differences in the expression modality. Non-native speakers used a greater variety of modal verb items, whereas native speakers used more adverbs with modalizing meanings.

As seen in aforementioned studies, the use of LBs did not focus on the comparison between native and non-native speakers until recent years. Comparison of different language backgrounds has a crucial role in applied linguistics since it gives insights in what and how to teach. Below are some studies done to compare LBs used by native and nonnative speakers.

Jalali et al. (2008) compared the use of LBs in two different corpora, one consisting of data from journals in applied linguistics and one consisting of data from master theses and doctoral dissertations written by Iranian post-graduate applied linguistics students. The results of the study revealed that not only are there relatively large differences between student genres and research articles in the types of bundles employed, but also there are some striking discrepancies in terms of frequency and function of the LBs used.

Takiand Jafarpour(2008)investigated the ways in which English and Persian academics express their position to discover the strategies used to bring readers to their writing. They compared the research articles written in English and Persian in two disciplines, Chemistry and Sociology. It was found that the writers of both disciplines,

considered the expression of stance and engagement markers in their writing important, but in Persian there seemed to be more cases of readers' involvement.

Chen & Baker (2010) analyzed a corpus of published academic texts and two corpora of student academic writing (one L1, the other L2). It was found that widest range of LBs was used in the published academic writing; whereas L2 student writing showed the smallest range. Furthermore, some high-frequency expressions in published texts, such as ***in the context of***, were underused in both student corpora, while the L2 student writers overused certain expressions, such as ***all over the world***, which native academics rarely used.

Wei & Lei (2011) compared the use of LBs by native speakers and Chinese EFL learners. The results revealed that the EFL learners used much more LBs than native speakers. They also used more passive structures and less anticipatory it structures than the professional writers.

CHAPTER III

METHODOLOGY

3.0. Introduction

The present chapter deals with the design of the study. First, the data collection tools and the participants of the study will be introduced. Then, the concordancing program, WordSmith, will be described. The explanation of the statistical analysis used in the study will follow. Finally, the last section will discuss thetaxonomies used for structural and functional analysis of the identified lexical bundles.

3.1. Data collection tools and participants

In order to answer the aforementioned research questions, International Corpus of Learner English (ICLE) and Louvain Corpus of Native English Essay (LOCNESS) are used to gather the data. Each corpus consists of argumentative essays. As Ebeling (2011) states, essay is a sub-category of academic writing. The academic setting of essay writing triggers certain recurrent word-combinations commonly used in academic writing in general; for instance *in order to, the fact that*.

Two learner corpora from ICLE*v*2, namely TICLE (Turkish International Corpus of Learner English) and JPICLE (Japanese International Corpus of Learner English) were selected as learner corpora. ICLE is chosen as the data source since it is the best known learner corpus with a collection of 200,000 words each from 14 different language backgrounds, such as German, Turkish, Finnish, Dutch and Spanish. Also, asAijmer (2009) points out, advanced learners make few morphosyntactic errors, but they use forms and structures in a non-native way.By comparing the essays produced within the ICLE with a native speaker norm, subtle features, such as overuse and underuse which account for the impression of near-nativeness of learners' essays, can be discovered. A strict criterion was set at the beginning of the ICLE Project considering the highly heterogeneous nature of learner language. According to the criterion, the learners are young adults (university undergraduates); they are in advanced proficiency level and are learning English as a Foreign Language (EFL) rather than as a Second Language (ESL). The language of the corpus consists of academic writing (mainly argumentative); 200,000 words per sub-corpus. Several other variables which

can influence learner productions were also included in the corpus; such as task variables (medium, genre, field, length, topic, and task setting) and learner variables (age, gender, mother tongue, region, other foreign languages, staying in English-speaking country, learning context, and proficiency level).As said previously, a wide range of essay topics are given to the participants. Top ten essay topics with the highest proportion are presented in Table 1.

Table 1

Top Ten Essay Topics in ICLE

Essay topics	**Number of essays**	**Country of origin**
Some people say that in our modern world, dominated by science, technology and industrialization, there is no longer a place for dreaming and imagination. What is your opinion?	491	29% Bulgarian
Most university degrees are theoretical and do not prepare students for the real world. They are therefore of very little value.	249	22% Turkish
Poverty is the cause of the HIV/AIDS epidemic in Africa.	243	100% Tswana
Marx once said that religion was the opium of the masses. If he was alive at the end of 20th century, he would replace religion with television.	237	19% Russian
The prison is outdated. No civilized country should punish its criminals; it should rehabilitate them.	176	32% Tswana
Discuss the advantages and disadvantages of banning smoking in restaurants.	156	100% Chinese
Discuss the advantages and disadvantages of using credit cards.	149	100% Chinese
Feminists have done more harm to the cause of women than good.	139	100% Russian
In the words of the old son “Money is the root of the evil”.	133	22% Russian
In his novel “Animal Farm”, George Orwell wrote “All men are equal: but some are more than others”. How true is this today?	127	39% Bulgarian

Source: Granger et al., 2009, p. 6-7.

Both TICLE and JPICLE consist of argumentative essays. This text type allows for discourse-oriented (cohesion, coherence, argumentative patterns, etc.) exploration as well as lexical and grammatical exploration (Aijmer, 2009).

TICLE contains 280 essays and 199,532 words. The essays were gathered from three institutions: University of Çukurova, the University of Mustafa Kemal and Mersin

University. The average length of essays was 713 words. The participants were undergraduate university students. The average age was 22,08. 81% of the participants were female and 195 were male students. In ICLE, each sub-corpus is divided into several batches with variable numbers of essays. A certain coding system is applied to identify them. Each batch of essays is given a 5-character code: First two is the code for nationality: e.g.JP for Japanese and TR for Turkish. These two letters are followed by the institution code: e.g. CU for Çukurova University. If there is more than one batch within an institution, it was given a single letter code. For instance, TRCU1 refers to Turkish Sub-corpus, Çukurova University, 1st batch. The following table summarizes the data gathered in TICLE.Text types are shown as: Argumentative (A), literary (L) or other (O).

Table 2

Turkish Sub-corpus

Batch	**Essay codes**	**Text types**			**Number of essays**	**Number of words**
		A	**L**	**O**		
TRCU1	TRCU1001-1177	177	0	0	177	128,297
TRKE2	TRKE2001-2072	72	0	0	72	50,228
TRME3	TRME3001-3031	31	0	0	31	21,007
TOTAL		280	0	0	280	199,532

Source: Granger et.al, 2009, p.37

JPICLE contains 366 argumentative essays and 198,131 words. Twenty-one institutions contributed to JPICLE project, some of which provided only one or two essays. The average length of essays was 542 words. The participants of the project were also undergraduate university students with an average age of 20,06. 73% of the participants were female and 27% was male. The table below summarizes the data gathered for JPICLE.

Table 3

Japanese Sub-corpus

Batch	Essay codes	Text types			Number of essays	Number of words
		A	L	O		
JPAI1	JPAI1001-1002	**2**	**0**	**0**	**2**	**1,451**
JPDO1	JPDO1001	**1**	**0**	**0**	**1**	**679**
JPFJ1	JPFJ1001	**1**	**0**	**0**	**1**	**622**
JPHI1	JPHI1001-1005	**5**	**0**	**0**	**5**	**2,887**
JPKO1	JPKO1 001-1020	**20**	**0**	**0**	**20**	**10,762**
JPKO2	JPKO2001-2031	**31**	**0**	**0**	**31**	**18,871**
JPKW1	JPKW1001-1002	**2**	**0**	**0**	**2**	**1,160**
JPKY1	JPKY1 001-1002	**2**	**0**	**0**	**2**	**1,206**
JPMI1	JPMI1001-1002	**2**	**0**	**0**	**2**	**1,123**
JPMJ1	JPMJ1001-1002	**2**	**0**	**0**	**2**	**1,057**
JPMU1	JPMU1001-1002	**2**	**0**	**0**	**2**	**972**
JPNH1	JPNH1001	**1**	**0**	**0**	**1**	**538**
JPOK1	JPOK1 001	**1**	**0**	**0**	**1**	**837**
JPRI1	JPRI1001-1002	**2**	**0**	**0**	**2**	**1,109**
JPSE1	JPSE1001	**1**	**0**	**0**	**1**	**834**
JPSH1	JPSH1001-1004	**4**	**0**	**0**	**4**	**2,422**
JPST1	JPST1001-1002	**2**	**0**	**0**	**2**	**1,248**
JPSW1	JPSW1001-1039	**39**	**0**	**0**	**39**	**18,338**
JPSW2	JPSW2001-2021	**21**	**0**	**0**	**21**	**11,691**
JPSW3	JPSW3001-3031	**31**	**0**	**0**	**31**	**16,532**
JPSW4	JPSW4001-4032	**31**	**0**	**0**	**31**	**16,735**
JPTF1	JPTF1001-1043	**43**	**0**	**0**	**43**	**23,011**
JPTK1	JPTK1001-1002	**2**	**0**	**0**	**2**	**987**
JPTM1	JPTM1001-1028	**28**	**0**	**0**	**28**	**16,793**
JPWA1	JPWA1001-1019	**19**	**0**	**0**	**19**	**9,433**
JPWA2	JPWA2001-2009	**9**	**0**	**0**	**9**	**4,291**
JPWA3	JPWA3001-3020	**20**	**0**	**0**	**20**	**10,097**
JPWA4	JPWA4001-4012	**12**	**0**	**0**	**12**	**7,257**
JPWA5	JPWA5001-5029	**29**	**0**	**0**	**29**	**14,649**
JPWA6	JPWA6001	**1**	**0**	**0**	**1**	**649**
Total		**366**	**0**	**0**	**366**	**198,131**

Source: Granger et.al, 2009, p.33

LOCNESS is the native speaker reference corpus compiled by the ICLE team at Louvain-la-Neuve. It is a corpus of native English essays made up of British pupils' A level essays (60,209 words), British university students' essays (95,695 words), and American university students' essays (149,501 words). In this dissertation, American university students' essays are used since the participants' age was similar to learner corpora and they were all argumentative essays. Therefore; the native speaker corpus in this study contains 168,400 words. Some of the universities participated in the Project were Marquetta University, University of Michigan, University of South Carolina and

Presbyterian College. The length of essays is around 500 words, similar to that of ICLE corpus.

LOCNESS was compiled by the Centre for English Corpus Linguistics at the Catholic University of Louvain, Belgium and it was made available for public use in 1998. The texts were gathered between 1991 and 1995. The corpus contains four components as essays of British A-level students, essays of British university students, argumentative essays of American students and literary-mixed essays of American students. Texts types of the corpus contain examination papers, timed essays and free essays. Reference tools were used in some timed and free essays whereas not used in examination papers. The age of participants ranges from 17 to 23 with a small number of older ages. There was a wide variety of essay topics, such as water pollution, nuclear power, sex, violence, gender roles, values and consequences of school interaction, and controversy in the classroom or prayer in schools.

Table 4

LOCNESS Selected Components

Institutions	**Codes**	**Text types**	**Number of essays**	**Number of words**
Marquette University	ICLE-US-MRQ	**Argumentative essay**	**46**	**54,285**
Indiana University, Indianapolis	ICLE-US-IND	**Argumentative essay**	**27**	**13,382**
Presbyterian College, South Carolina	ICLE-US-PRB	**Argumentative essay**	**6**	**12,447**
University of South Carolina	ICLE-US-SCU	**Argumentative essay**	**53**	**52,885**
University of Michigan	ICLE-US-MICH	**Argumentative essay**	**43**	**16,502**
Total			**175**	**149,501**

3.2. Data Analysis

In this study, both descriptive and quantitative type of research design was conducted in order to answer the research questions mentioned in the previous chapter. Table 5below shows the steps followed in conducting the study.

Table 5

Research Type and Stages of the Study

Stages	**Process**	**Research type**
Stage1	Automatic generation of frequency lists	Descriptive analysis
Stage2	Selection of meaningful LBs manually	Corpus-based approach
Stage3	Application of statistical analyses across corpora	Quantitative corpus analysis
Stage4	Analysis of LBs structurally and functionally	Descriptive analysis
Stage5	Comparison of LBs among 3 corpora	Interpretative

As can be seen in table 5 above, first of all, all the four-word word combinations are determined in LOCNESS, TICLE and JPICLE respectively by using the WordSmith analysis tool. Among these occurrences of four-word combinations, stance lexical bundles were selected manually. After selecting them, the identified LBs in each corpus were evaluated in terms of their structures and functions. The calculation of the frequencies of these structural and functional types followed this step. Finally, the use of stance lexical bundles by the native speakers and EFL learners was compared.

Since the focus of this proposal is stance lexical bundles, while doing the analysis, the 1st person singular and plural pronouns are included and considered as personal stance; whereas, 2nd and 3rd person singular and plural pronounsare excluded since as Biber (2006; 91) points out 'Although such structures expresssome kind of attitude or evaluation, they do not necessarily reflect the personal stance of the speaker/writer'.In this study, the 2nd person is only included in the analysis when it was an attitudinal/modality stance; such as *if you want to*, and *it is necessary to*.

Also, when analyzing the data, bundles including abbreviation are combined with the ones with no abbreviation so that they do not stand as different word groups. For instance, *I do not want* and *I don't want*, and *it is better to* and *it's better to* are combined and counted as one bundle. Similarly, bundles that are similar structurally and functionally are combined and counted as one. For instance, I *think the*, *I think it*, *I think this*and *I think we* are all SP+VP structure and they are all impersonal and uncertain. Therefore; these bundles are combined and counted as one impersonal/uncertain and SP+VP bundle. Similarly, the verbs 'is' and 'are' are combined and the bundles counted as one, such as *is/are very important*. In the analysis, the frequencies of each bundle are summed up.

3.2.1. Cut-off Point

As stated previously, Biber et al. (1999, p. 990) define a lexical bundles as ‘recurrent expressions’ regardless of their idiomaticity, and regardless of their structural status’.They are sequences of words that commonly go together. A word combination has to meet a set of defining criteria to be counted as a bundle. Firstly, they must occur frequently in a register. They are simply the mostfrequently occurring sequences of words. The frequency cut-off point of the bundles varies from study to study. Biber (2006) concluded that to be a lexical bundle, a four-word expression had to recur ten times per million words and had to appear in more than five texts. On the other hand, Biber, Conrad, & Cortes (2004) set the criteria as a lexical bundle had to occur forty times in a one-million word corpus; whereas, Cortes (2004) decided to set the cut-off point at twenty times in one million words. These higher cut-off points were chosen to be more conservative in the frequency of these expressions and to ensure that the object of analysis in these studies consisted of unit expressions that were used in extremely high frequencies. In addition to frequency, lexical bundles must be used in at least five different texts (10%). This prevents focusing on idiosyncratic uses by the authors of the texts in the corpus under consideration. Since high cut-off points are chosen to be more conservative, in this study, the cut-off point is determined as 3. Also, the bundles which are seen in at least 3 different texts are chosen to make the analysis more accurate. Therefore, the bundles that are not used by 3 different participants are not included.

As stated previously, the cut-off point in the analysis is 3; therefore, bundles that are used 3 times in three different texts were included. Some of the bundles, such as *it is said* (28) and *it is said that* (25), and *want to say* (30) and *I want to say* (20), are both included in the analysis since *it is said* and *it is said that* are considered different structures; the former SP+VP and the latter DC fragment, and both of them are used more than three times in three different texts. Similarly, *I want to say* is used more than three different texts and it is important that the subject ‘I’ denotes personal stance.

Three and four-word lexical bundles are analyzed in this dissertation. Although ‘many four-word bundles hold three-word bundles in their structures’ (Cortes, 2004, p.401)and they are more common and present a wider range of structures and functions (Hyland, 2008b), three-word lexical bundles are also included since many they have a crucial role in the analysis. For example, the bundles *the fact that* and *I think that* were used very frequently in all corpora; therefore, it is important to analyze them.

3.2.2. WordSmith

In corpus linguistics, computer software and corpus tools are essential for researchers in reaching empirical conclusions and analyzing the data. In this proposal, the concordance program WordSmith v.5, created by Scott (2008), is used in determining the lexical bundles. As Scott (2010) states, WordSmith Tools is an integrated suite of programs for looking at how words behave in texts. WordSmith Tools can be used to produce *frequency* lists, to run *concordance* searches and calculate collocations for particular words, and to *find keywords* in a text and examine their distribution (Baker, Hardie and McEnery, 2006). WordSmith can be considered an efficient tool to identify word combinations after meeting the previously-established cut-off points for frequency. In this study, WordList tool and Concordance tool were used. The main page of WordSmith can be seen in Figure 4 below.

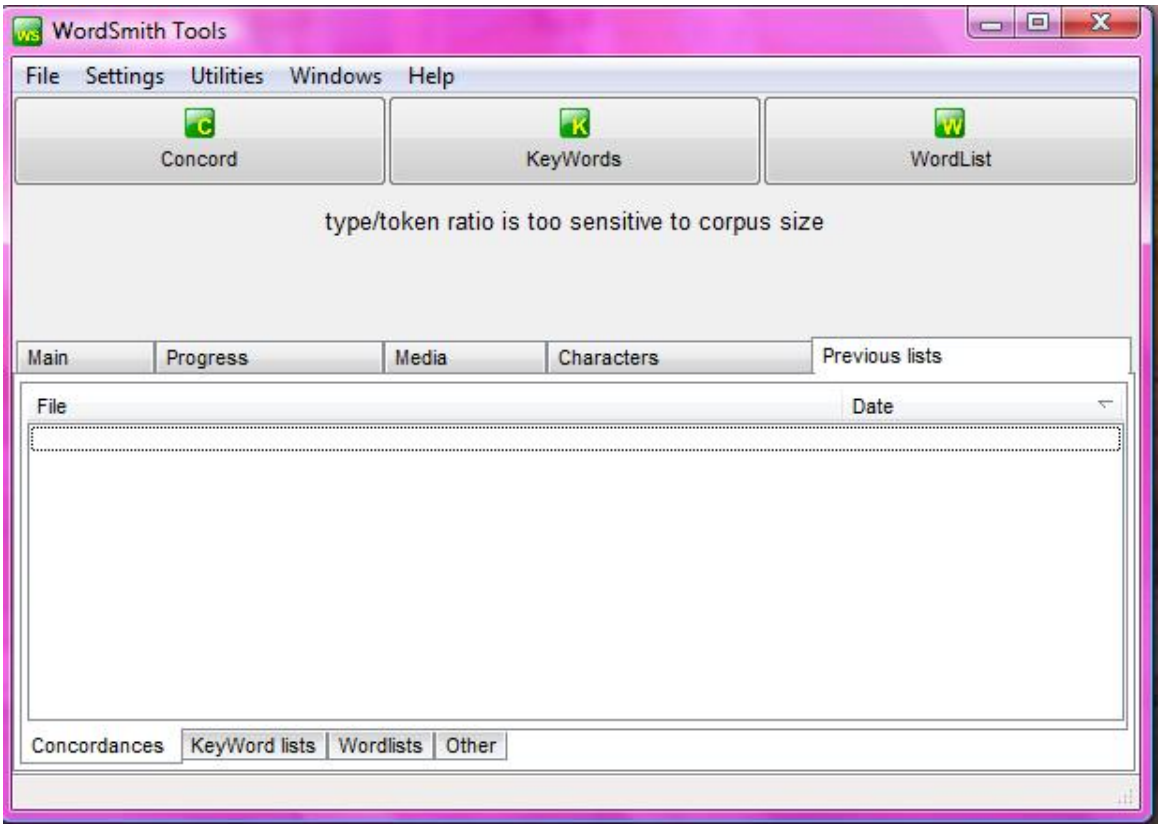


Figure 4. A Screenshot of WordSmith main page

With the WordList tool one can see a list of all the words or word-clusters in a text which can be set out in alphabetical or frequency order. It helps the language researcher identify the common words in a corpus, information which is useful for

example when determining which lexical items to teach and which to ignore, or when the materials writer is attempting to ensure that new vocabulary is met more than once in a textbook (Scott, 2001).

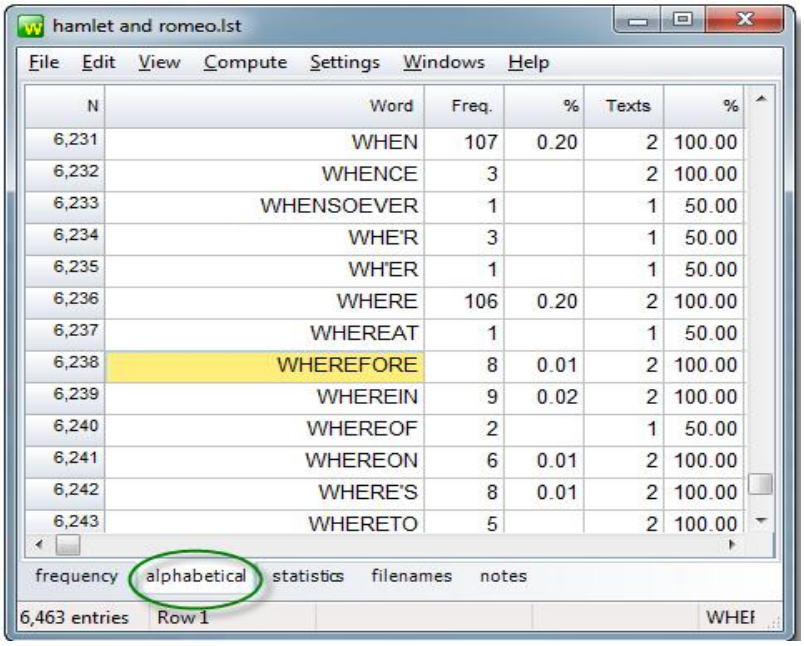


Figure 5. A Screenshot of WordList page in WordSmith Tools

Source: http://www.lexically.net/wordsmith/step_by_step_English6/index.html

With the Concord tool, one can see any word or phrase in context; it helps to see with what kind of word(s) surround the word or phrase that is studied. It lists the occurrence of given word or phrase in a corpus (Scott, 2001). The point of a concordance is to be able to see lots of examples of a word or phrase, in their contexts.



Figure 6. A Screenshot of Concord Tool in WordSmith Tools

Source: http://www.lexically.net/wordsmith/step_by_step_English6/index.html

3.2.3. Type-token Ratio

To find out the lexical variety, type-token ratio, referred to as T/t hereafter, is calculated. T/t ratio is a measure of vocabulary variation within a written text or a person's speech and it is a helpful measure of lexical variety within a text (Williamson, 2013). T/tis obtained by dividing the type count by the token count, which is always <= 1. A high T/t indicates a high degree of lexical variation while a low T/t indicates the opposite. T/t can also be expressed as a percentage, multiplying the ratio by 100 (Thomas, 2005). T/t ratio of lexical bundles represents the percentage of each bundle within all words in a corpus, that is, the number of bundles that falls into per 100 words. The T/t ratio is calculated according to the following formula:

Type-token ratio = (number of types/number of tokens) * 100

The number of words in a text is often referred to as the number of 'tokens'. However, several of these tokens are repeatedin corpora (Williamson, 2009). For example, the sample Table 6 below shows all the tokens with their frequency occurrence. In this sample, the token ***again*** occurs two times and the token ***we*** occurs six times. If we count the number of words, we get a total of 87. That total of the 87 tokens, there are 62 so-called 'types'. The relationship between the number of types and the number of tokens for this sample is:(62/87) * 100 = 71.3%. It means that 71.3 of the words are from a certain type of group in 100 words. The more types there are in comparison to the number of tokens, then the more varied is the vocabulary; that is, lexical variety is greater.

Table 6

Sample Table for Type-token Ratio

rank	word	freq	rank	word	freq	rank	word	freq	rank	word	freq
1	we	6	17	asleep	1	33	impressions	1	49	seem	1
2	and	5	18	at	1	34	instant	1	50	so	1
3	them	5	19	beliefs	1	35	is	1	51	the	1
4	are	3	20	brief	1	36	just	1	52	then	1
5	can	3	21	but	1	37	later	1	53	things	1
6	they	3	22	call	1	38	metaphor	1	54	thinking	1
7	to	3	23	coming	1	39	mull	1	55	this	1
8	again	2	24	concepts	1	40	notions	1	56	thoughts	1
9	as	2	25	described	1	41	occasions	1	57	times	1
10	in	2	26	endure	1	42	opinions	1	58	values	1
11	on	2	27	fall	1	43	other	1	59	variously	1
12	a	1	28	going	1	44	over	1	60	views	1
13	act	1	29	handle	1	45	perceptions	1	61	well	1
14	all	1	30	have	1	46	put	1	62	what	1
15	an	1	31	however	1	47	refer	1	TOTAL		87
16	aside	1	32	ideas	1	48	return	1			

Source: Williamson, 2009

3.2.4. Log Likelihood

In addition to T/t ratio, frequency distribution of the corpora is also tested. To do this, Log Likelihood (LL) analysis, which is sometimes called G-square or G-score, is used. LL test considers word frequencies weighted over two different corpora. It measures higher or lower frequencies than expected. In the analysis, LOCNESS frequency results are compared with TICLE and JPICLE (L1 vs. L2 comparison). In order to see the condition among learner groups, overall lexical bundle frequency of TICLE and JPICLE are also compared (L2 vs. L2 comparison).

LL compares the observed and expected values for two datasets. Observed valuesarethe actual frequencies extracted from corpora, andexpected values are the frequencies that one would expect if no factor other than chance were affecting the values. The greater the difference between the observed and the expected values, the less likely it is that the difference has arose by chance.Log-likelihood is calculated by constructing a contingency (frequency) table.

	Corpus 1	Corpus 2	Total
Frequency of word	A	B	a+b
Frequency of other words	c-a	d-b	c+d-a-b
Total	C	D	c+d

Figure 7. Contingency Table

Source: Rayson & Garside, 2000

The value 'c' corresponds to the number of words in corpus one, and'd' corresponds to the number of words in corpus two (N values). The values 'a' and 'b' are called the observed values (O), the expected values (E) are calculated according to the following formula:

$$E_i = \frac{N_i \sum_i O_i}{\sum_i N_i}$$

In the formula, N1 = c, and N2 = d. So, for this word, E1 = c*(a+b) / (c+d) and E2 = d*(a+b) / (c+d). The calculation for the expected values takes account of the size of the two corpora, so we do not need to normalize the figures before applying the formula. We can then calculate the log-likelihood value according to this formula:

$$-2\ln \lambda = 2\sum_i O_i \ln\left(\frac{O_i}{E_i}\right)$$

This equates to calculating log-likelihood G2 as follows: G2 = 2*((a*ln (a/E1)) + (b*ln (b/E2))).Contingency table as presented in Figure 7 has two rows and two columns and has 1 degree of freedom. In LL, critical values with 1 df are 3.84, 6.64 and 10.83 for the significance levels of 0.05, 0.01 and 0.001. For example, 28.841 (1 df) is greater than 10.83 for the significance level 0.001. Therefore, we are more than 99.99 percent confident that difference in the frequencies observed in two corpora is statistically significant. If the LL results are greater than 6.64; the difference between two corpora happening by chance is less than 1%. If LL value is 3.84 or more, theprobability of it happening by chance is less than 5%.

3.2.5. Analyzing Functions and Structures of Stance Lexical Bundles

To answer the second and third research questions (2. What are the structural and functional features of the stance lexical bundles used by native English speakers, and Turkish and Japanese EFL learners? 3. To what extend do the stance lexical bundles used by the Turkish and Japanese EFL learners have in common with the ones used by the native speakers?), functional and structural categories of lexical bundles are determined manually. When determining them, Biber et al's (2004) functional and structural categories of lexical bundles are used since it has been widely relied on in the studies on lexical bundles in the field. However, while determining the epistemic stance bundles, a revised version of their classification was used.

As Biber et al. (2004) states, epistemic stance bundles comment on the knowledge status of the information (certain, uncertain, probable/possible), such as *I don't know if,* and *I don't think so.* There are two sub-categories of epistemic stance bundles: Personal and impersonal. Personal epistemic bundles can express certainty and uncertainty but most of them express only uncertainty. *I don't know what, I don't think so,* and *I think it was* are examples of personal epistemic bundles. Impersonal epistemic bundles express certainty; *are more likely to,* and *the fact that the* are considered in this group. However, a slight change has been applied to this model by adding two more categories since while analyzing the data, it was found that there were many personal certain and impersonal uncertain lexical bundles. Therefore, epistemic stance bundles were categorized as:

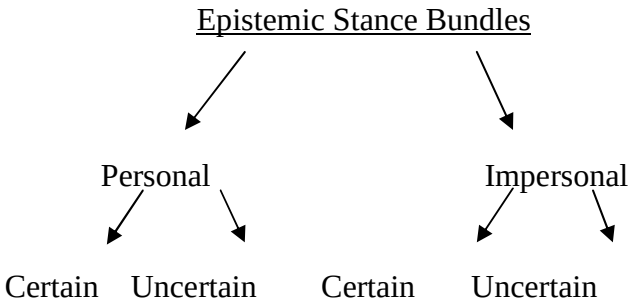


Figure 8. Categories of epistemic stance bundles

When analyzing the data, it was found that some chunks expressing 'evaluation' are fairly used. These chunks indicate what the writer thinks about a specific subject. They are formed by using an adjective and they make an assessment on a particular topic. Since they make an assessment on a particular topic, they give a clear indication of the writers' thoughts on that specific topic. Therefore, these chunks also show personal stance and they can be counted as a new category of stance lexical bundles. This new category is added to the Attitudinal/Modality category. Hence, the new attitudinal/modality category consists of these groups: Desire, obligation/direction, intention/prediction, ability, and evaluation.Biber (2006, p. 92) categorizes common lexico-grammatical features used for the stance analyses and he lists some stance adjectives used in evaluation. Some of the adjectives he mentions in this category are appropriate, good/bad, better/worse, essential, important, interesting and necessary. While determining the evaluation stance lexical bundles, this list is used as a basis. Some of the lexical bundles determined in the evaluation category are: *It is difficult to,*

is the best, is very easy to, and *is very interesting.*The functional categories used in this dissertation can be seen in Figure 4 below.

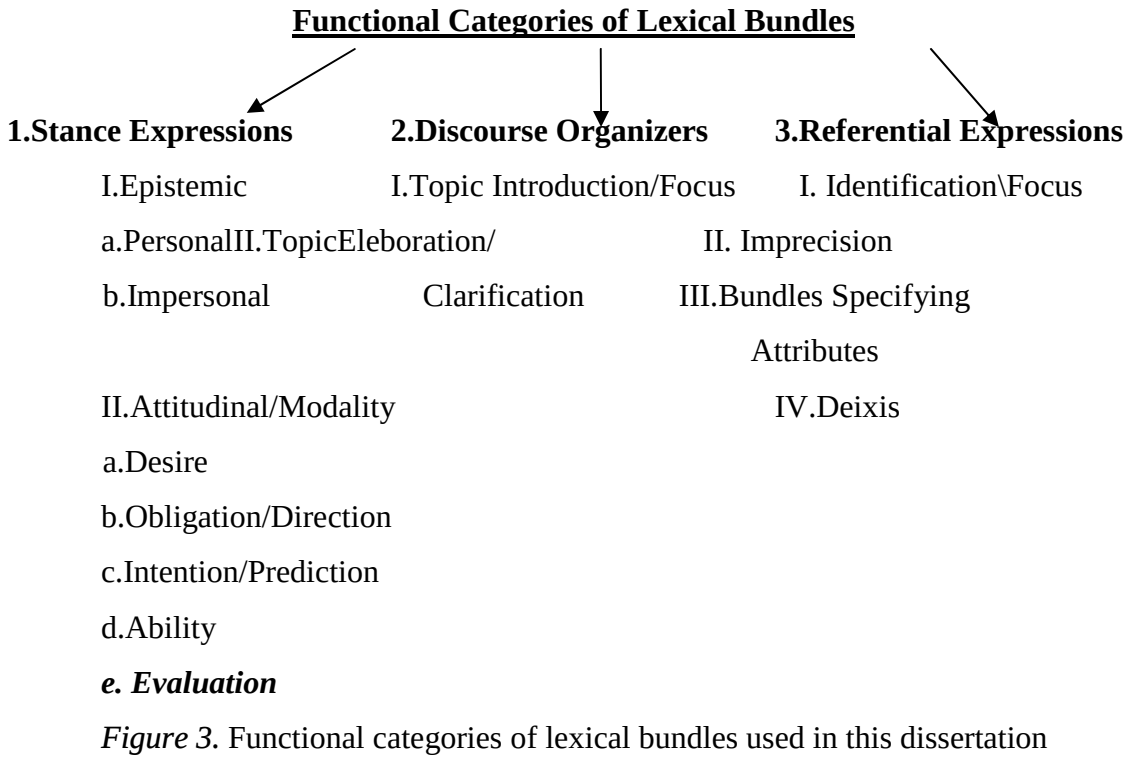


Figure 3. Functional categories of lexical bundles used in this dissertation

CHAPTER IV

RESULTS AND DISCUSSIONS

4.0. Introduction

This chapter deals with the results of the stance lexical bundles which are analyzed with descriptive statistics through software programs. Firstly, descriptive results of the SLBs used in each corpus are given. Secondly, functions of stance lexical bundles used in three corpora are presented and their use is compared. Thirdly, the structures of stance lexical bundles are presented and a comparison of their use is done. Finally, an overall comparison of SLBs is done across three corpora by means of L1 and L2 and L2 and L2 comparison.

4.1. Overall Results of the Study

In the first step of the analysis, the stance lexical bundles used in corpora were gathered and calculated regardless of their category. Each stance lexical bundles from nine types was identified over LOCNESS, TICLE, and JPICLE. By concordancing via software, all instances of stance lexical bundles found over corpora, then the number of each instance were calculated. Finally, a total frequency ofeach stance lexical bundle was obtained. The frequency calculation was made in order todetermine the proportion of these bundles in L1 and L2 corpora; thus thefrequency results might be compared to each other. The overall descriptive results of the stance lexical bundles used in the texts are given in the Table 7 below. The list of all stance lexical bundles can be seen in Appendices A, B and C.

Table 7

Frequencies and Percentages of Stance Lexical Bundles

Corpora	Total number of words	Total number of SLBs	% of SLBs
TICLE	199,532	288	0,14
LOCNESS	149,501	184	0,12
JPICLE	198,131	357	0,18
Total	547,164	829	0,44

As it is seen in Table 7 above,the total number of the words analyzed is 547,164 and 0,44% of these words consist of stance lexical bundles. As the results show, native speakers use lexical bundles least.Japanese EFL learners use SLBs more frequently than the native speakers and Turkish EFL learners. Japanese EFL learners used 357 LBs, whereas native speakers used 184 and Turkish EFL learners used 288. The results are consistent with Wei and Lei (2011) and Jalali et al. (2008) that nonnative speakers use much more lexical bundles than native speakers. These results do not coincide withChen and Baker (2010) for in their study, they found that the widest range of LBs was used by native speakers and the smallest range was used by nonnative learners.

The number of lexical bundles found in this dissertation is higher than the numbers found in previous studies. For instance, Hyland (2008a) found 629 bundles in a corpus of nearly 3.5 million, Wei and Lei (2011) found 241 bundles in a 2.5 million corpus. In these studies, they analyzed the 4-word lexical bundles, since as Wei and Lei state, 3-word bundles are extremely common. Therefore, one of the reasons of the high number of lexical bundles found in this study is both three and four-word lexical bundles are included. Also, as mentioned in the previous chapter, a new category (evaluation) is added to Biber's classification. The highest number of lexical bundles is in this category in the Japanese corpus. Also, the cut-off point for the lexical bundles was set 3 since the size of the corpus was not very big. As seen in the results, the frequency of the bundles is various; therefore, in order to determine whether there is an overuse or underuse, Log Likelihood analysis is done. The results of this analysis are as follows.

4.1.1. Log Likelihood Analysis Results

In order to test the frequency distribution of the corpora, Log Likelihood (LL) analysis is done. LL test considers word frequencies weighted over two different corpora. It measures higher or lower frequencies than expected. LOCNESS frequency results are compared with orderly TICLE and JPICLE (L1 vs. L2 comparison). In order to see the condition among learner groups, overall lexical bundle frequency of TICLE and JPICLE are also compared (L2 vs. L2 comparison).

Initial LL analysis was made between LOCNESSand TICLE corpora to test the overall frequency difference occurred between them. Table 8 shows the LL ratio of the lexical bundles in comparison with TICLE and LOCNESS.

Table 8

LL Ratio of TICLE and LOCNESS

	O1 TICLE	**%1**	**O2 LOCNESS**	**%2**	**LL ratio**
Lexical Bundles	288	0,14	184	0,12	+2,88

O1 is observed frequency in Corpus 1
O2 is observed frequency in Corpus 2
%1 and %2 values show relative frequencies in the texts.
+ indicates overuse in O1 relative to O2,
- indicates underuse in O1 relative to O2

In the table, O1 and O2 refer to overall frequency of lexical bundles observed in TICLE and LOCNESS. %1 value includes the relative frequency of lexical bundles in the texts, i.e., 0,14 relative frequency means there are approximately 0,14 lexical bundles fall into every 100 words in TICLE. Similarly, 0,12 bundles fall into every 100 words in LOCNESS. According to the result, LL ratio measurement indicates an overuse in TICLE with an +2,88 LL value.This means that there is a significant difference between two corpora in terms of lexical bundle frequency ($p < 0.05$) and the overuse in TICLE relative to LOCNESS has been approved by LL calculation.

Next comparison is made between JPICLE and LOCNESS and the result of this comparison is presented in Table 9below.

Table 9

LL Ratio of JPICLE and LOCNESS

	O1 JPICLE	**%1**	**O2 LOCNESS**	**%2**	**LL**
Lexical Bundles	357	0,18	184	0,12	+18,28

O1 is observed frequency in Corpus 1
O2 is observed frequency in Corpus 2
%1 and %2 values show relative frequencies in the texts.
+ indicates overuse in O1 relative to O2,
- indicates underuse in O1 relative to O2

As expected from the frequency difference between JPICLE and LOCNESS, LL value between the two corpora revealed an overuse of lexical bundles. Japanese EFL learners used 0,18lexical bundles in every 100 words, whereas native speakers used 0,12lexical bundles in every 100 words. The LL ratio between these two corpora is found + 18,28. In the LL analysis, the higher the LL value means the more significant

overuse; therefore we can conclude that the overuse of lexical items in JPICLE is more than the use of TICLE.

Last LL ratio comparison is made between two learner groups, namely Japanese and Turkish EFL learners. The result of this comparison is as follows:

Table 10

LL Ratio of JPICLE and TICLE

	O1 JPICLE	**%1**	**O2**	**%2 TICLE**	**LL**
Lexical Bundles	357	0,18	288	0,14	+7,89

O1 is observed frequency in Corpus 1
O2 is observed frequency in Corpus 2
%1 and %2 values show relative frequencies in the texts.
+ indicates overuse in O1 relative to O2,
- indicates underuse in O1 relative to O2

The LL ratio of JPICLE and TICLE show that Japanese EFL learners overused stance lexical bundles with a+7,89 LL ratio. The results of LL shows that compared to native speakers,both EFL learner groups overused lexical bundles; however,since the language item under consideration is an English language structure, the usage of stance lexical bundles in LOCNESS is accepted as the accurate usage.

4.1.2. Type-token Ratio Results

Table 11shows the type-token ratio results of the three corpora, respectively LOCNESS, TICLE and JPICLE. T/t ratio shows the lexical variety within a text. T/t ratio of lexical bundles represents the percentage of each bundle within all words in the corpora, that is, the number of bundles that falls into per 100 words.

Table 11

Overall Comparative Frequency Distribution of Stance Lexical Bundles in General in Three Corpora

	Tokens	**Types**	**T/t Ratio (<=1 / %)**
LOCNESS (L1)	1056	184	0,17 (17%)
TICLE (L2)	2519	288	0, 11 (11%)
JPICLE (L2)	8394	357	0,04 (4%)

n= raw frequency of epistemic bundles
T/t ratio= Type/token ratio; percentage of number of stance lexical bundles (types) in total of words (tokens) in each corpus

As Table 11 shows, the lowest number of lexical bundles in total belongs to LOCNESS (184). Yet, when the T/t ratio of LBs is compared, it was seen that the T/t ratio of LBs used in LOCNESS is 0,17 (17%); whereas, it is 0,11 (11%) in TICLE, and 0,04 (4 %) in JPICLE.This means that native speakers used more varied stance lexical bundles; 17 of their every 100 words were a stance lexical bundle. Turkish learners used 11 of their every 100 words as a stance lexical bundle and Japanese learners used 4. The percentage of T/t ratios is similar in LOCNESS and TICLE corpora, but it is lower than these two groups in JPICLE.Therefore, the distribution of epistemic and attitudinal/modality stance lexical bundles are also compared. The results are presented in the Table 12 below.

Table 12

Overall Comparative Frequency Distribution of Epistemic and Attitudinal/Modality Stance Lexical Bundles in Three Corpora

	Epistemic SLBs			**Attitudinal/Modality SLBs**		
	Tokens	**Types**	**T/t Ratio (<=1 / %)**	**Tokens**	**Types**	**T/t Ratio (<=1 / %)**
LOCNESS (L1)	772	118	0,15 (15%)	284	66	0,23 (23%)
TICLE (L2)	1425	184	0,13 (13%)	1094	104	0,09 (9%)
JPICLE (L2)	1496	170	0,11 (11%)	6898	187	0,03 (3%)

n= raw frequency of epistemic bundles
T/t ratio= Type/token ratio; percentage of number of epistemic and attitudinal/modality stance lexical bundles (types) in total of words (tokens) in each corpus

As Table 12 presents, the percentage of the use of epistemic stance lexical bundles is more than the attitudinal/modality stance lexical bundles in L2 learner groups, namely Turkish and Japanese EFL learners. Turkish EFL learners used 13 of their every 100 words as epistemic SLBs and Japanese EFL learners used 11 of their every 100 words as epistemic stance lexical bundles. On the contrary, native speakers used attitudinal/modality SLBs more than epistemic SLBs.The difference between the use of epistemic and attitudinal/modality lexical bundles is the highest in Japanese EFL learners group. They used epistemic bundles 11%; whereas attitudinal/modality bundles 3%. The reason for this could be that Japanese EFL learners used evaluation lexical bundles excessively compared to other bundles. When the most frequent 40 lexical bundles are compared (Table 13), it is seen that Japanese learners used some of the bundles far more frequently than the EFL learners. For instance, *I think that* was used

279 times by the Japanese EFL learners; whereas it was used 29 times by the Turkish EFL learners and 9 times by the native speakers. Therefore, although the number of LBs is more in JPICLE, the lexical variety is lower. Conversely, native speakers used the fewest number of lexical bundles; however lexical variety is the highest (17%) in this group. That is, native speakers use less but various bundles; whereas EFL learners use more bundles, but they are not various; that is, similar or the same bundles are used in many instances. The results are consistent with Chen (2012) that in her study, she also found that nonnative speakers use restricted range of epistemic devices. In order to see the differences between the three corpora, the most common 40 bundles are compared.

4.1.3. Forty Most Common StanceLexical Bundles

Table 13 presents the most common 40 stance lexical bundles used in LOCNESS, TICLE and JPICLE. The bundles that are used in all of the three corpora are written Italic and bold.

Table 13

Most Common 40 Stance Lexical Bundles

NATIVE		**TURKISH**		**JAPANESE**	
Stance Bundles	**Fre.**	**Stance Bundles**	**Fre.**	**Stance Bundles**	**Fre.**
THE FACT THAT	71	***I THINK+N (THE/IT/THIS/etc.)***	88	***I THINK+N (IT/THE/THIS,etc.)***	309
BE ABLE TO	50	IN MY OPINION	70	I THINK THAT	279
HAVE TO BE	23	I WANT TO	59	I WANT TO	185
ARE/IS GOING TO	22	BE ABLE TO	53	WE HAVE TO	108
I FEEL THAT	20	***I BELIEVE THAT***	43	WE NEED TO	94
DO NOT WANT	18	IT CAN BE	39	***IS/ARE VERY IMPORTANT***	68
BE ALLOWED TO	15	DO NOT WANT	36	I DO NOT THINK	67
NEED TO BE	15	IT WILL BE	32	IT IS IMPORTANT	48
WILL CONTINUE TO	14	***IS/ARE VERY IMPORTANT***	31	***DON'T HAVE TO***	43
THE OPPONENTS OF	13	***THE FACT THAT***	31	WE WANT TO	40
I BELIEVE THAT	12	THERE WILL BE	31	***I BELIEVE THAT***	38
THERE WOULD BE	12	WE CAN SEE	30	I WOULD LIKE TO	38
WOULD HAVE TO	12	***I THINK THAT***	29	IT WILL BE	38
IT IS IMPORTANT	11	DO NOT KNOW	27	I DON'T WANT	34
WOULD LIKE TO	11	DO NOT THINK	30	ARE ABLE TO	32
IS/ARE FORCED TO	10	THERE SHOULD BE	24	IT IS GOOD	32
WOULD HAVE BEEN	10	WE CAN SAY	23	MOST IMPORTANT THING	31
ARGUE THAT THE	9	WE SEE THAT	23	WILL BE ABLE TO	31
DO NOT HAVE TO	9	WE HAVE TO	22	HAVE TO BE	30
I THINK THAT	9	IS THE MOST IMPORTANT	21	IS VERY DIFFICULT	30

(Table 13 Continued)

I THINK+N (IT/THE)	9	IS/ARE GOING TO	19	IT IS DIFFICULT	30
IN MY OPINION	9	IT MEANS THAT	19	WANT TO SAY	30
IT IS OBVIOUS	9	***DO NOT HAVE TO***	18	I DON'T KNOW	28
THERE WILL BE	9	CAN BE DONE	17	IS VERY USEFUL	28
CAN BE SEEN	8	I CAN SAY	17	IT IS SAID	28
I HAVE SEEN	8	IS/ARE NOT IMPORTANT	17	I NEED TO	25
PEOPLE FEEL THAT	8	THERE IS NO NEED	17	IT IS NATURAL	25
SHOULD BE ALLOWED TO	8	PEOPLE THINK THAT	16	IT IS SAID THAT	25
WILL BE ABLE TO	8	WE KNOW THAT	16	IS GOOD FOR	24
WILL HAVE TO	8	IN FACT THERE/THIS/THE	15	IS THE BEST	23
WOULD BE ABLE TO	8	IT IS NECESSARY	15	IT IS NECESSARY	23
ARE WILLING TO	7	SEEMS TO BE	15	IT MAY BE	23
CONSIDERED TO BE	7	THE IMPORTANT THING IS	15	***THE FACT THAT***	23
I WOULD LIKE TO	7	WE CAN SAY THAT	15	IT IS VERY IMPORTANT	22
IS/ARE ALLOWED TO	7	ACCORDING TO ME	14	I AM GOING TO	21
IS VERY IMPORTANT	7	DO NOT UNDERSTAND	13	IT IS TRUE	21
IT IS POSSIBLE	7	IT IS BETTER	13	IT MIGHT BE	21
MAY HAVE BEEN	7	IT IS TRUE	13	WILL NOT BE	21
ONE CAN SEE	7	SEEN AS A/AN	13	DON'T NEED TO	20
SHOULD BE ABLE TO	7	THIS MEANS THAT	13	I WANT TO SAY	20

To understand the difference between the native speakers and EFL learners, it is necessary to compare which epistemic stance bundles are used by them. As it is seen in Table 13; *thefact that, do not have to, I think that,*and *is/are very important* are the most common lexical bundles used by both the native speakers and EFL learner groups. When the frequency of these bundles is compared, it is found that*I think+n*is by far the most frequently used stance lexical bundle by the Japanese EFL learners. The most frequently used lexical bundle by the Turkish EFL learners is *I think+n*(88), whereas the most frequently used lexical bundle by the native speakers is*the fact that*(71). *The fact that* is considered an impersonal certain epistemic stance bundle and it is used 31 times by the Turkish EFL learners, 23 times by the Japanese learners. Unlike native speakers, EFL learners used personal uncertain epistemic stance bundles the most, which are*in my opinion,I think+n* (it/the/this, etc.) and *I think that. In my opinion* is used 70 times by the Turkish EFL learners; on the other hand, it was used 9 times by the native speakers and 20 times by the Japanese EFL learners. *I think that* is used excessively by the Japanese EFL learners; they used it 279 times. However, native speakers used it only 9 times and Turkish EFL learners used it 29 times. Another striking finding is Japanese EFL learners' use of the verb 'think'. They used it immensely; *I think+n* is used 309 times and as stated previously, *I think that* is used 279

times. Turkish EFL learners used *I think+n* 88 times, which was the most frequently used bundle in this group. Unlike Japanese and Turkish EFL learners, native speakers used *I think that* and *I think+n* only 9 times. As discussed in T/t ratio previously, Japanese EFL learners used some of the bundles, such as *I think that* and *I want to*, excessively; therefore, although they used more bundles than the native and Turkish EFL learner groups, the variety of the bundles Japanese EFL learners used is less than the variety of the bundles the other groups used.

It is also seen that Japanese EFL learners tend to use first person singular and plural pronouns more often than the native speakers and Turkish EFL learners. 'I' is used 11 times and 'we' is used 3 times by the Japanese EFL learners in the most common 40 lexical bundles. Native speakers used 'I' five times and they did not use 'we'. The use of 'I' and 'we' by the Turkish EFL learners remains in between the native speakers and Japanese EFL learners. They used 'I' three times and 'we' five times. Surprisingly, Japanese speakers also used the expletive 'it' more frequently than the native speakers. Japanese is a pro-drop language; therefore it is surprising that Japanese native speakers used the expletive it more than the other groups. It is also surprising that although they used the dummy subject it more frequently than the other groups, they did not use the expletive 'there'. Turkish is also a pro-drop language, but Turkish native speakers did not use the expletive 'it' as often as Japanese native speakers.English native speakers used expletive pronouns 'it' and 'there' twice. They used the first person singular six times and they did not use the first person plural pronoun.The use of first person subject pronouns and expletives by the Turkish EFL learners remains in between the native speakers and Japanese EFL learners. It can be concluded from these results that the use of epistemic devices in general by the Turkish EFL learners is closer to native speaker use. Hyland (2005) suggests that contrary to what is often thought, academic writing is not an impersonal monologue. Use of self-mention is an important stance marker, but as seen in the results; native speakers these expressions less frequently than EFL learners.on the other hand, EFL learners used them more often. This finding is inconsistent with Taki&Jafarpour(2012) that native speakers use self-mention more than the learners, namely Persian group. Another important finding the table shows is that Japanese EFL learners use uncertainty verbs 'think' and 'believe' immensely. Conversely, the verb choice of native speakers is varied: believe, argue, seem, think, and feel, and the most frequently used epistemic bundle by the native speakers is 'the fact that', which shows certainty. It can be concluded that Japanese

learners used uncertainty bundles more often compared to native speakers and Turkish EFL learners. The reason for this could be the difference between the cultures. As stated in Kamimura and Oi (1998), Japanese argumentative style is more subjective rather than factual or objective, and there is more weight on evoking empathy in the audience, rather than more emphasis on persuading the audience.

4.2. Functions of Stance Lexical Bundles

To answer the second and third research questions (2. What are the structural and functional features of the stance lexical bundles used by native English speakers, and Turkish and Japanese EFL learners? 3. To what extend do the stance lexical bundles used by the Turkish and Japanese EFL learners have in common with the ones used by the native speakers?), functional and structural categories of lexical bundles are determined. The functions and structures of lexical bundles are determined by using Biber et al.'s (2004) categories as stated in the previous chapter. The summary of the results are shown in Chart 1 below.

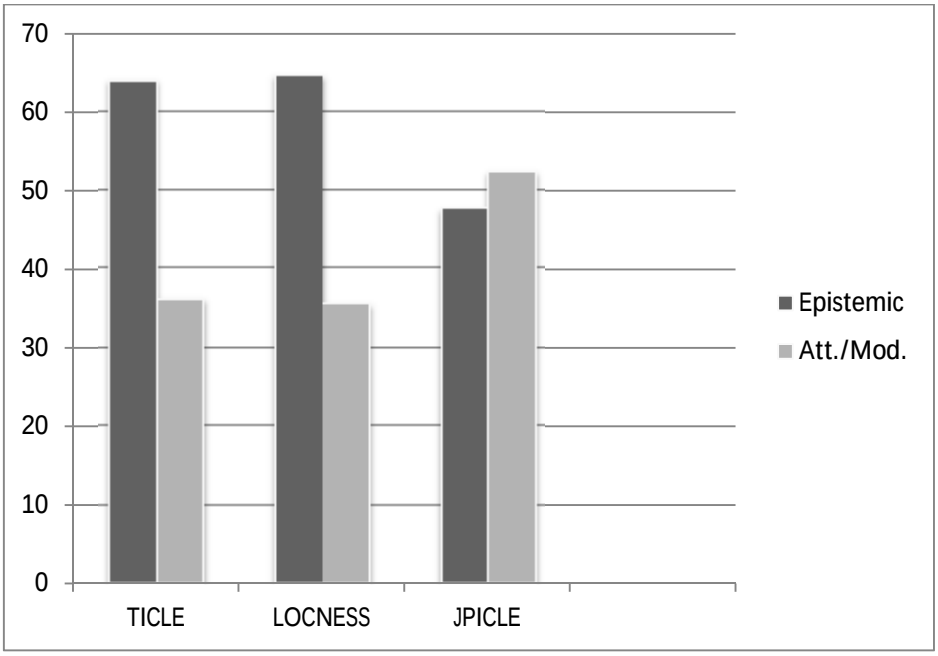


Chart 1. Functions of lexical bundles in general

As seen in Chart 1, epistemic bundles are used more frequently in TICLE (63,89%) and LOCNESS (64,14%); whereas attitudinal/modality bundles are used more frequently in JPICLE (52,37%). When the use of epistemic stance bundles and

attitudinal/modality stance bundles are compared in all groups, it is found that epistemic bundles are used most frequently by the native speakers (64,14%). The results are somewhat consistent with Chen (2012) that in her study she found that there is a great similarity in the total number of epistemic devices used by native and non-native (Chinese) speakers. In the current study, the percentage of the use of epistemic devices by the native speakers and Turkish EFL learners is very close to each other; however the percentage of Japanese EFL learners is lower than the percentage of native speakers.Also, as found in T/t ratio results, both Turkish and Japanese EFL learners used more restricted range of devices to express epistemic stance, which is more typical of academic prose. Native speakers used epistemic bundles slightly more than the Turkish EFL learners; whereas they used themrelatively more than the Japanese EFL learners. The reason why Japanese EFL learners used attitudinal/modality lexical bundles more is that they used evaluation bundles excessively. Looking at each functional category would give more insights on the use stance lexical bundles. Table 14 below summarizes the functions of lexical bundles used in LOCNESS, TICLE, and JPICLE more in depth.

Table 14

Functions of Stance Lexical Bundles in Detail

FUNCTIONS OF STANCE LEXICAL BUNDLES

Corpora		**TICLE**		**LOCNESS**		**JPICLE**	
Frequency & Percentages		**Freq.**	%	**Freq.**	%	**Freq.**	%
EPISTEMIC BUNDLES	P/C	**41**	**14,24**	**14**	7,6	**37**	10,36
	P/UC	**18**	6,25	**7**	3,8	**47**	**13,17**
	IM/C	**65**	22,57	**47**	**25,54**	**48**	13,45
	IM/UC	**60**	20,83	**50**	**27,2**	**38**	10,64
ATTITUDIONAL/ MODALITY BUNDLES	Desire	**11**	3,82	**7**	3,8	**21**	**5,88**
	Obligation/Direction	**29**	10,07	**29**	**15,76**	**47**	13,17
	Intention/Prediction	**11**	3,82	**9**	**4,89**	**15**	4,20
	Ability	**3**	1,04	**2**	1,09	**7**	**1,96**
	Evaluation	**50**	17,36	**19**	10,32	**97**	**27,17**
	Total	**288**	100	**184**	100	**357**	100

As can be seen in Table 14 above, epistemic stance bundles are frequently usedin all groups. Both native speakers and Turkish and Japanese EFL learners used impersonal stance bundles pervasively. That is, compared to personal stance bundles, impersonal stance bundles are used more often. The total number of impersonal stance bundles used by the native speakers is 52,74%. Similarly, both Japanese (24,09%) and Turkish (43,4%) EFL learners used impersonal stance, such as ***are more likely to***, more often than the personal stance, such as *I think that*. Among all lexical devices used by native speakers, 64,14% of them were epistemic devices, most of which (52,74%) shows impersonality. Similarly, 43,4% out of 63,89% of the Turkish data consists of impersonal epistemic devices, and 24,09% out of 47,62% of Japanese data consists of impersonal epistemic devices. When the use of certainty and uncertainty devices is compared, it is seen that native and Turkish EFL learners used certainty devices more than uncertainty, whereas the number of certainty and uncertainty devices was the same in Japanese native speakers. Turkish EFL learners used 78 (27,06%) uncertainty devices and 106 (36,81%) certainty devices. Native speakers used 57 (31%) uncertainty and 61 (33,14%) certainty devices. Japanese EFL learners used the same amount of certainty and uncertainty devices: 85 (23,81%).The use of certainty/uncertainty devices by the native speakers is slightly more than each other (2,14%), whereas the use of certainty/uncertainty devices by the Turkish EFL learnersis a little more visible (9,75%). In order to make better inferences about the data,personal certainty/uncertainty and impersonal certainty/uncertainty should also be compared.

When personal certainty and uncertainty is compared, it is seen that personal certain lexical bundles, such as ***I know that, we have seen*** are more common than personal uncertain lexical bundles, such as ***I think that, it seems to me*** in LOCNESS and TICLE; on the contrary, personal uncertain lexical bundles are used more frequently in JPICLE.Native speakers used more personal certain lexical bundles. It can be concluded that when self mention personal words are used, they preferred using words with certainty; whereas words with uncertainty were preferred in impersonal lexical bundles. Contrary to what Biber et al. (2004) state, personal certain lexical bundles are used more frequently in these two corpora, namely TICLE and LOCNESS. Turkish EFL learners(14,24%) used personal certain lexical bundles most frequently, butJapanese EFL learners (13,17%) used personal uncertain lexical bundles most frequently. Although Japanese EFL learners used personal uncertain lexical bundles more frequently, their use of these devices is slightly more than personal certain lexical

devices (10,36%). Hyland (2005) says that contrary to what is often thought, academic writing is not an impersonal monologue. He (2002 a, p. 1110) also states that 'Self-mention constitutes a central pragmatic feature of academic discourse since it contributes not only to the writer's construction of a text, but also of a rhetorical self. The authorial pronoun is a significant means of promoting a competent scholarly identity and gaining acceptance for one's ideas. Since use of self-mention is an important stance marker, it should be wise to look at the use of self mention words; such as the subject pronouns 'I' and 'we', object pronouns'me' and 'us', and possessive pronouns'my' and 'our'. The table below summarizes the self mention words used in LOCNESS, TICLE and JPICLE.

Table 15

Use of Self Mention in Corpora

Self mention words	**TICLE (freq.)**	**LOCNESS (freq.)**	**JPICLE (freq.)**
'I' (I think that, I've seen that, I cannot say, etc.)	388	107	1383
'WE' (We don't know, we can understand, we must think, etc.)	254	32	416
'**MY'** (in my opinion, in my view, my opinion is that)	70	9	26
'ME' (according to me, it seems to me, to me that)	14	7	24
Total	726	155	1849

As can be seen, the subject pronouns 'I' and 'we', and the object pronoun 'me' are used most frequently by the Japanese EFL learners. They used the subject pronouns excessively. The possessive pronoun 'my' is used most frequently by the Turkish EFL learners. Although Turkish EFL learners used 'my' most frequently, it was used in only one lexical bundle, which is *in my opinion.* Similarly, they used 'me' in only one lexical bundle: *according to me.* Both native speakers and Japanese EFL learners used *to me that* and *it seems to me*. *To me that* was used 14 times by the Japanese EFL learners and 4 times by the native speakers. *It seems to me* was used 10 times by the Japanese EFL learners and 3 times by the native speakers. It is surprising that Japanese EFL learners used self mentioning more than the native speakers and Turkish EFL learners since as Ohta (1991) states, first person pronoun is largely unacceptable because collective identity is more valued than individuality in Asian cultures. However, the results are consistent with Ishikawa (2009). It was also found that Japanese speakers of English

overused the pronoun *I* and the expression *I think,* which shows a subjective point of view. The second group which used self mention frequently was Turkish EFL learners. Uysal (2012) compared the argument preferences of Turkish people when writing in Turkish (L1) and when writing in English (L2). She compared the indirectness markers and found that there is a similar use of indirectness markers in L1 and L2, specifically point of view of distancing devices, such as 'I think' and 'I believe that'. When writing in L2, the participants usedthese devices a little less than they do in their L1. She concluded that when writing in English, Turkish speakers are influenced by their L1. That is to say, these expressions are also used in Turkish and Turkish native speakers use these devices when writing in English (L2) as well.

When impersonal certainty and uncertainty is compared, it was found that Turkish and Japanese EFL learners used more impersonal certain lexical bundles than impersonal uncertain lexical bundles; however native speakers used more impersonal uncertain lexical bundles. Turkish EFL learners used more certainty words in both personal and impersonal bundles. However, Japanese EFL learners used more certainty words in impersonal lexical bundles; whereas they used more uncertainty words with personal lexical bundles.

Attitudinal/modality stance lexical bundles were also compared. As Table 14 presents, Japanese EFL learners used desire, ability and evaluation lexical bundles more frequently than the native speakers and Turkish EFL learners. Compared to other attitudinal/modality lexical bundles, obligation/direction and evaluation lexical bundles were used most frequently by all the participants. Native speakers used obligation/direction and intention/prediction lexical bundles more frequently than the EFL learners. The results are consistent with Tenuta et al. (2012) that in their study they also found that native speakers used modals that denote intention/prediction and obligation/direction more frequently than the nonnative speakers.

As seen in the results, Japanese EFL learners used evaluation lexical bundles, such as *it is not good, is easy to, it is wonderful, do not like* and *it is so hard* excessively. 27,17% of the stance bundles they used were evaluative. Kamimura and Oi (1998) compared the argumentative essays written by native English speakers and Japanese EFL learners and found that while more rational appeals are used in Western rhetoric, more affective appeals are used by Japanese speakers, giving more weight on evoking empathy on the reader. They use more softening and emotional devices, such as 'I think that' and 'sad'. This may be reason that Japanese EFL learners used more

evaluative lexical bundles; whereas English native speakers used them the least. The use of evaluative lexical bundles by the Turkish EFL learners was in between the native speakers and Japanese EFL learners.In order to understand the use of stance lexical bundles, the structures used in each corpus should also be compared.

4.3. Structures of Stance Lexical Bundles

Structures of stance lexical bundles are determined by using Biber et al.'s (2004) categories as stated in the previous chapter. The summary of the results are shown in Chart 2 below.

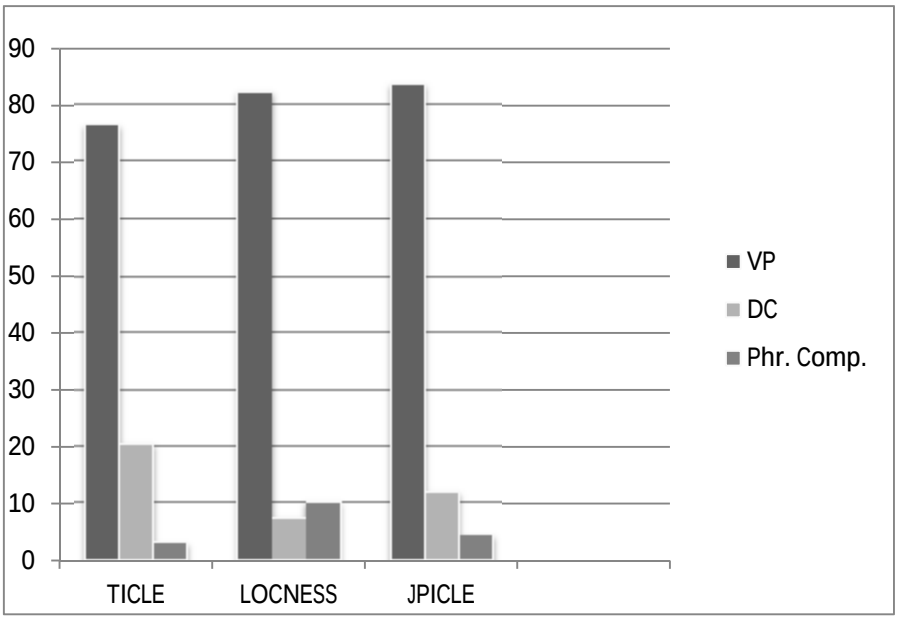


Chart 2. Structures of stance lexical bundles in general

As Chart 2 above shows, verb-phrase structures were used most frequently in all corpora. 83,47% of the bundles were verb-phrase structures in JPICLE. 83,15% of bundles constitute verb phrase structures in LOCNESS and 76.39% in TICLE. Dependant clauses follow verb-phrase structures in TICLE (20,49%) and JPICLE (12,05%). Phrasal components were used the least in TICLE (3,12%) and JPICLE (4,48%); whereas 9,78% of bundles was phrasal components in the native speaker data.This result is inconsistent with the literature stating that noun phrase structures are immensely used (Strunkyte and Jurkunaite, 2008; Bal, 2010; Wei and Lei 2011). However, it is consistent with Hernandez (2013). Structures of lexical bundles should

be analyzed in-depth in order to compare them effectively. Table 16 presents the structures of lexical bundles used in the three corpora.

Table16

Structures of Stance Lexical Bundles in Detail

STRUCTURES OF STANCE LEXICAL BUNDLES							
Corpora		**TICLE**		**LOCNESS**		**JPICLE**	
Frequency & Percentages		**Freq.**	%	**Freq.**	%	**Freq.**	%
VP FRAGMENTS	**SP+VP**	**109**	37,85	**75**	40,76	**177**	**49,58**
	VP	**110**	38,19	**77**	**41,85**	**119**	33,33
	Q	**1**	0,35	**1**	0,54	**2**	0,56
DEPENDENT CLAUSE	**DC**	**59**	**20,49**	**13**	7,07	**43**	12,05
PHRASAL COMPONENTS	**NP+Post Mod.**	**7**	2,43	**13**	**7,07**	**12**	3,36
	PP+Emb.	**2**	0,69	**5**	2,71	**4**	1,12
	TOTAL	**288**	100	**184**	100	**357**	100

As stated previously, the most commonly used structure was VP fragments in all corpora. Dependant clause structures follow VP fragments in TICLE and JPICLE, whereas phrasal components follow it in LOCNESS. VP fragments consist of a subject pronoun followed by a verb phrase, such as *I'm not going to, I want to*; a verb phrase, such as *is very important, don't have to*, or a question, such as *do we have to, do we think that.*Native speakers and Japanese EFL learners used more SP+VP structures than VP; whereas Turkish EFL learners used VP structures slightly more than SP+VP. Question forms are very rarely used in all corpora. Dependent clause structures are reasonably used by the EFL learners (Turkish 20,49%, Japanese 12,05%). Native speakers used DCs 7,07%. The highest number of phrasal components (9,78%) was used by native speakers. There is an insignificant difference in the use of VPs between

Turkish and Japanese EFL learners. The results are consistent with Hernandez (2013) that all three groups used VP fragments most frequently. Both in this dissertation and Hernandez's study, native speakers used more phrasal components than dependent clause structures. In the learner data, it is seen that dependent clause is used more frequently than phrasal components. Conversely, phrasal components are used more frequently in Hernandez (2013). The reason why Turkish and Japanese EFL learners use more DC structures could be related to their mother tongue. As stated in Allen (2009), there is a similar structure, like *it can be said that*, used in Japanese. Likewise, similar structures are used in Turkish.

The findings do not coincide with Strunkyte and Jurkunaite (2008) that in their study they found that NP components are used most frequently followed by VP and DC fragments. Similarly, Bal (2010) found that Turkish scholars use more PP and NP fragments than VP fragments. Wei and Lei (2011) found a similar result that native speakers used NPs more often; whereas Chinese EFL learners preferred VPs.

The structures of 40 most common lexical bundles are also compared in detail. The comparison can be seen in Table 17 below, which gives a broad idea about the structures used in native and nonnative corpora.

Table 17

Structures of 40 Most Common LBs in Detail

Structures		TICLE	LOCNESS	JPICLE
	NP based	1	2	2
NP/PP	PP based	3	1	-
	Copula be+NP/Adj	3	2	5
VP	VP with active verb	5	1	1
	Passive verb	2	4	2
	VP+that clause	-	1	-
	Modals+verbs	2	17	4
	SP+VP	9	4	15
SP+VP	Dependent clause	6	3	2
	There is/are	3	2	-
	Anticipatory it+VP/Adj Pr.	6	3	9
TOTAL		40	40	40

When the structures of the 40 most common lexical bundles were compared, the results showed that prepositional phrases were not used in JPICLE; whereas a small number of PP was used in TICLE and LOCNESS. VP+that clause was used only in LOCNESS: *argue that the*. The pronoun 'there' is not used by Japanese EFL learners.

Copula be+NP/Adj phrase was mostly used in JPICLE, the reason for this is that evaluative lexical bundles were more frequently used by Japanese EFL learners and the adjectives used in this group were evaluative. Similarly, anticipatory it+VP/Adjphrase was mostly used by Japanese EFL learners. Again, the reason for that is the excessive use of evaluative lexical bundles by the Japanese. Native speakers used more passive structures than EFL learners;which coincides with Wei and Lei (2011). These results are not consistent with Hyland 2008a and Wei& Lei, 2011that in their studies, they found that learners used less anticipatory it and more passive structures. This could be because teaching materials and professors in China suggest learners to use devices of impersonality, such as passives. The conflicting results may due to cultural differences.

4.4. Comparison of the Three Corpora

The last research question (Do the differences/similarities regarding the use of lexical bundles stand from the learners' first language or is this an interlanguage property?) tried make a comparison between L1 vs. L2 and also L2 vs. L2 in order to investigate the (common) interlanguage properties. In order to answer this question, frequency of lexical bundles identified in LOCNESS and TICLE and JPICLE (L1 and L2), and also TICLE and JPICLE (L2 and L2) were compared.The obtained results were evaluated by means of frequency analysis and then compared through Log Likelihood analysis for the statistical confidence of frequency comparison. Then, theuse of self-mention words was also compared.

In the analysis, it was found that EFL learners used more stance bundles (Turkish EFL learners 288, Japanese EFL learners 357) than native speakers (184). According to the frequency and LL analysis, there is certain overuse of lexical bundles in EFL learners' essays when compared tonative speaker group (L1 vs. L2). That is, both Turkish and Japanese EFL learners used lexical bundles more frequently than the native speakers. Japanese EFL learners used lexical bundles far more frequently than the native speakers with a +18,28 LL ratio.When the LL ratio of JPICLE and TICLE was compared (L2 vs. L2), an overuse of lexical bundles by the Japanese EFL learners was observed. That is to say, Japanese EFL learners used lexical bundles more frequently than both groups, namely native speakers and Turkish EFL learners.The overall results of LL ratio showed that lexical bundle usage by the Turkish learners remains somewhere between the native speakers' and the Japanese learners' usage.The

general tendency of overuse of stance lexical bundles by the EFL learners might be due to interlanguage properties which tend learners to determine certain ways of using stance lexical bundles. Stance lexical bundles have not been determined in Turkish and Japanese languages yet; therefore, it is not possible to make a totally accurate interpretation whether the (over)use of stance lexical bundles is related to the learners' native language (L1 transfer), cultural writing styles or cultural writing habits.However, the detailed analysis of the functions, structures and the use of self mention in the corpora give us a comprehensivepicture of how to express stance.

When the functions of stance lexical bundles are compared, the results revealed that epistemic stance bundles are used more frequently than attitudinal/modality stance bundles by the native speakers and Turkish EFL learners; however the use of epistemic and attitudinal/modality stance markers is just the opposite in Japanese EFL learners. When personal and impersonal bundles are compared, the results showed that impersonal stance bundles are used more frequently than personal bundles in all groups. When the certainty and uncertainty are compared, it was seen that native and Turkish EFL learners used certainty devices more than uncertainty; whereas the number of certainty and uncertainty devices was the same in Japanese native speakers. This is somewhat in linewithAğçam (2014) that in her study, nonnatives (Turkishand Spanish) usedmorecertaintydevicesthannativespeakers.When the im/personal un/certainty is compared, it is seen that native speakers use impersonal uncertain and personal certain lexical devices more (than personal uncertain and impersonal certain). Turkish EFL learners used personal certainty devices more than personal uncertaintydevices; however, their use of impersonal certainty devices was more than personal certainty devices. Similarly, Japanese EFL learners used impersonal certainty devices more than impersonal uncertainty devices. Yet, they used personal certainty devices less than personal uncertainty. It can be said concluded that native speakers are more confident in expressing their point of view.

When the use of self mention words and most common 40 stance lexical bundles are looked at, it is found that self mention words (I, we, my, and me) were used most frequently by the Japanese EFL learners (1849). Turkish EFL learners follow them with a total number of 729. Native speakers used these words least (155). The analysis of the 40 most common lexical bundles showed that ***I think+n*** was the most frequently used stance lexical bundle by the Turkish and Japanese EFL learners. However; native speakers used ***the fact that***as the most frequent SLB. It can be said that EFL learners

used more subjective devices; on the other hand native speakers used more objective devices when expressing stance.

When the structures of stance lexical bundles are compared, it was found that verb-phrase structures were used most frequently in all corpora, which is in line with Hernandez (2013). Dependent clauses follow VP structures in EFL data. In LOCNESS, native speaker corpus, phrasal components follow VP structures. This could be related to the participants mother tongue that as Allen (2009) states, similar structures; such as *it can be said that*, are used in Japanese. Likewise, these structures are also used in Turkish. It is surprising that both Turkish and Japanese EFL learners used anticipatory it+VP/Adj pronouns more than the native speakers. Both Turkish and Japanese are pro-drop languages, but both groups used these structures more frequently than the natives. However, it is seen that native speakers used more passive structures. These results are inconsistent with Wei and Lei (2011) and Hyland (2008a) that in their studies, they found that learners used less anticipatory it and more passive structures. The conflicting results may due to cultural differences and the more use of evaluative lexical bundles by the EFL learners.

The next chapter will assess the study and the results found taking the research questions and theoretical approaches into consideration. Furthermore, the implications and recommendations for language learning and teaching will be presented with the limitations of the study and suggestions for future research.

CHAPTER V

CONCLUSION

5.0. Introduction

This chapter summarizes the findings of this dissertation briefly. Subsequent to the summary and review of the research questions, the limitations of the study are given. Finally, some implications and recommendations for future research are presented.

5.1. Summary of the Study

Lexical bundles have a crucial role in both academic prose and spoken discourse. Succinctly put, fluent linguistic production depends on them; they are important in successful interaction and gaining the control of a new language. As mentioned previously, although they have a very crucial role in interaction, very few studies have been done comparing the use of lexical bundles by native and nonnative speakers of English. Regrettably, none of these studies focus on the comparison of native speakers and nonnative speakers with different language backgrounds. Also, these studies do not analyze the functions and structures of these bundles in depth. Therefore, this study aimed to find answers to the following research questions:

Research Question 1: What are the most common stance lexical bundles found in argumentative essays written by native English speakers, and Turkish and Japanese EFL learners?

The first aim of this study was to determine which stance lexical bundles are used by the native speakers and Turkish and Japanese EFL learners and to determine which lexical bundles were used most frequently.The results of the analysis showed that 184 stance lexical bundles were used by the native speakers, 288 stance lexical bundles were used by the Turkish EFL learners and 357 stance lexical bundles were used by the Japanese EFL learners. Yet, when the type token ratio of the bundles was compared, it was seen that native speakers use less but various stance lexical bundles. The results showed that 17 % of the words native speakers used were stance lexical bundles.

Whereas; 11 % of the words Turkish EFL learners used were stance lexical bundles, and 4 % of the words Japanese EFL learners used were stance lexical bundles. That is 17of the words in 100 words used by native speakers were stance lexical bundles, 11 of the words in 100 words used by Turkish EFL learners were stance lexical bundles, and 4 of the words in 100 words used by Japanese EFL learners were stance lexical bundles. Therefore, it can be concluded that although the number of lexical bundles is higher in Japanese and Turkish data, lexical variety in these groups is lower than the native data. In other words, native speakers used the fewest number of lexical bundles; however the lexical variety was the highest. When the most frequent stance lexical bundles were looked at, it was seen that the most common lexical bundles used by both the native speakers and EFL learners were *the fact that, I think that, I think+n, I believe that, do not have to,* and *is/are very important*.

Research Question 2: What are the structural and functional features of the stance lexical bundles used by native English speakers, and Turkish and Japanese EFL learners? Research Question 3: To what extent are the stance lexical bundles used by the Turkish and Japanese EFL learners similar to the ones used by native speakers?

When the functions of stance lexical bundles are compared, it was seen that epistemic stance bundles were used more frequently than attitudinal/modality stance bundles by the native speakers and Turkish EFL learners. On the contrary, Japanese EFL learners used attitudinal/modality stance bundles more often than the epistemic bundles.Epistemic stance bundles are used most frequently by the native speakers, which is consistent with the findings of Chen (2012) that nonnative speakers use more restricted range of devices to express epistemic stance.

When each functional category is compared in detail, it was found that impersonal stance bundles, such as *are more likely to*are used more often than the personal stance bundles, such as *I think that* in all groups. When the use of certainty/uncertainty devices was compared, it was seen that Turkish EFL learners and native speakers used certainty devices more than uncertainty; whereas the number of certainty and uncertainty devices was the same in Japanese EFL learners.

When the use of attitudinal/modality stance bundles are compared, it was found that obligation/direction and evaluation lexical bundles were used more frequently than the other groups of attitudinal/modality lexical bundles by both the native speakers and

EFL learners. It was also seen that native speakers used obligation/direction and intention/prediction lexical bundles more frequently than the EFL learners. Japanese EFL learners used desire, ability and evaluation lexical bundles more frequently than the native speakers and Turkish EFL learners. These results are consistent with Tenuta et al. (2012) that EFL learners used modals denting ability more of than native speakers and native speakers used modals that denote intention/prediction and obligation/direction more frequently than nonnative speakers.

The comparison of the structures of stance lexical bundles showed that verb phrases were by far the most frequent structures used by all the groups of speakers. Dependant clauses follow verb-phrase structures in TICLE and JPICLE; phrasal components were used the least. On the contrary, phrasal components were used more than the dependant clauses by the native speakers. SP+VP structures were used more frequently by the native speakers and Japanese EFL learners, whereas, VP fragments were slightly more than SP+VP structures by the Turkish EFL learners. Question forms are very rarely used by all the groups. NP fragments are used more than the PP fragments by both the native speakers and EFL learners.

Research Question 4: Do the differences/similarities regarding the use of lexical bundles stand from the learners' first language or is this an interlanguage property?

The last research question tried make a comparison between L1 vs. L2 and also L2 vs. L2 in order to investigate the (common) interlanguage properties. In order to answer this question, frequency of lexical bundles identified in LOCNESS and TICLE and JPICLE (L1 and L2), and also TICLE and JPICLE (L2 and L2) were compared.The obtained results were evaluated by means of frequency analysis and then compared through Log Likelihood analysis.A detailed comparison of the functions and structures of stance lexical bundles are also done.

In the analysis, it was found that EFL learners used more stance bundles than native speakers. According to the frequency and LL analysis, there is certain overuse of lexical bundles in EFL learners' essays when compared to native speaker group (L1 vs. L2). That is, both Turkish and Japanese EFL learners used lexical bundles more frequently than the native speakers. When the LL ratio of JPICLE and TICLE was compared (L2 vs. L2), an overuse of lexical bundles by the Japanese EFL learners was observed. That is to say, Japanese EFL learners used lexical bundles more frequently

than both groups. The overall results of LL ratio showed that lexical bundle usage by the Turkish learners remains somewhere between the native speakers' and the Japanese learners' usage. The general tendency of overuse of stance lexical bundles by the EFL learners might be due to interlanguage properties which tend learners to determine certain ways of using stance lexical bundles. Stance lexical bundles have not been determined in Turkish and Japanese languages yet; therefore, it is not possible to make a totally accurate interpretation whether the (over)use of stance lexical bundles is related to the learners' native language (L1 transfer), cultural writing styles or cultural writing habits. However, the detailed analysis of the functions, structures and the use of self mention in the corpora give us a comprehensive picture of how to express stance.

When the functions of stance lexical bundles are compared, the results revealed that epistemic stance bundles are used more frequently than attitudinal/modality stance bundles by the native speakers and Turkish EFL learners; however the use of epistemic and attitudinal/modality stance markers is just the opposite in Japanese EFL learners. When personal and impersonal bundles are compared, the results showed that impersonal stance bundles are used more frequently than personal bundles in all groups. When the certainty and uncertainty are compared, it was seen that native and Turkish EFL learners used certainty devices more than uncertainty; whereas the number of certainty and uncertainty devices was the same in Japanese native speakers. When the im/personal un/certainty is compared, it is seen that native speakers use impersonal uncertain and personal certain lexical devices more (than personal uncertain and impersonal certain). Turkish EFL learners used personal certainty devices more than personal uncertainty devices; however, their use of impersonal certainty devices was more than personal certainty devices. Similarly, Japanese EFL learners used impersonal certainty devices more than impersonal uncertainty devices. Yet, they used personal certainty devices less than personal uncertainty. It can be said concluded that native speakers are more confident in expressing their point of view.

When the use of self mention words and most common 40 stance lexical bundles are looked at, it is found that self mention words (I, we, my, and me) were used most frequently by the Japanese EFL learners. The use of these words is in between in the Turkish EFL learners and the native speakers used them least. The analysis of the 40 most common lexical bundles also showed that EFL learners used more subjective devices; on the other hand native speakers used more objective devices when expressing stance.

When the structures of stance lexical bundles are compared, it was found that verb-phrase structures were used most frequently in all corpora, which is in line with Hernandez (2013). Dependent clauses follow VP structures in EFL data. In LOCNESS, native speaker corpus, phrasal components follow VP structures. This could be related to the participants mother tongue that as Allen (2009) states, similar structures; such as *it can be said that*, are used in Japanese. Likewise, these structures are also used in Turkish. It is surprising that both Turkish and Japanese EFL learners used anticipatory it+VP/Adj pronouns more than the native speakers. Both Turkish and Japanese are pro-drop languages, but both groups used these structures more frequently than the natives. However, it is seen that native speakers used more passive structures. These results are inconsistent with Wei and Lei (2011) and Hyland (2008a) that in their studies, they found that learners used less anticipatory it and more passive structures. The conflicting results may due to cultural differences and the more use of evaluative lexical bundles by the EFL learners.

5.2. Implications and Recommendations for Language Learning and Teaching

Hyland (2008 a) states that formulaic language is difficult to acquire and that failing to use it correctly identifies learners as outsiders. As Jafarpour, Hashemian and Alipour (2013) point out, language teachers should attract their learners' attention to collocations, the co-occurrence of words, since grammar and individual words are not sufficient in language proficiency. Different academic discourse relies on different lexical clusters (Hyland, 2008 b), therefore lexical bundles specific to different genres should be taught and students should be made aware of these differences and the importance of these expressions. Cortes (2006) conducted a study in which she explicitly taught lexical bundles in a writing intensive history class. She found that students' awareness and interest in these expressions increased after the implementation. Neely & Cortes (2009) also conducted a study and they concluded that in order to design lessons and materials for the ESL/EFL classrooms, corpus-based research findings in conjunction with publicly available corpora and concordance programs can be used. Bundles should be taught by presenting the complete spectrum of their functions in context by analyzing these functions in discourse similar to the one they encounter daily in their academic lives. As Sinclair (2001) says, corpus can be used not only in finding the accurate description of structure and reliable models of usage,but

it can also be used inseeing how words and phrases are actually translated, what the essentials in a syllabus are, and what the characteristic errors of learners are. Small corpora can be put together quickly for a classroom job or anindividual need. Therefore, language teachers should be made aware of corpus-based teaching and how they can use corpora in their classrooms to make students more aware of these structures and more native like speakers. Finally, textbooks and other teaching materials should be prepared considering these frequently used structures.However, as Flowerdew (2001) suggests, pedagogic materials and teaching methods should be corpus based, not corpus bound.

Although the effectiveness of using corpora is seen in the studies conducted, using corpora in language teaching is challenging. While preparing corpus-based activities, there are several things the teachers shouldtake into consideration: the language and ability level of their students, the age of their students, and the cultural and educational background of their students. Saeedand Waly (2009)suggest some classroom activities that can be used in different language levels. They suggest that verb pairing games, computer cloze activity and screen shot analysis can be used with beginner level students. With intermediate level students, contextual analysis, cloze activities, spot the error activity, express yourself and multiple choice activities can be done and with the advanced level students, concordance analysis and pragmatics: situation analysis activities can be done.As Wu (?) points out, when using corpora in language teaching, the learners should be highly motivated and autonomous. They should have the ability to analyze and interpret linguistic data; therefore discovery learning is more suitable for advanced level students.

How to use corpora in classroom should be taught to preservice teachers. When applying corpus linguistics to language teacher education, certain problems can be faced. These should be kept in mind and solutions to them should be found: Where to get relevant corpus for teacher trainees, how to design corpus-aided activities matching the ongoing daily teaching goals, how to make huge corpus data manageable in classroom teaching, how to evaluate the effectiveness of teaching, how to guide trainees to make corpus analysis, and how to select and edit corpus example to meet learners' proficiency levels.(He, 2008)As stated previously, although it is challenging to use corpora in language teaching, it is a useful method in seeing the accurate usage of structural patterns and language functions at the word, phrase, clause and discourse level. It also shows what is common, what is possible and what is changing in the

language system (Smith, Butler, Griffith and Kritonis, 2007). Therefore; using corpora should not be neglected and it should be a part of classroom teaching.

5.3. Limitations of the Study

This study is restricted to only stance lexical bundles used by native speakers and two nonnative learner groups coming from different language backgrounds. Analyzing discourse organizers and referential expressions would give more insights on the use of lexical bundles. Moreover, only Turkish and Japanese EFL learners' use of LBs is compared. Comparing the use of LBs used by learners coming from other language backgrounds would give more ideas on these devices. Another limitation of this study was the size of the corpora. A small scale of corpus can present only limited findings; therefore, the results of this study cannot be generalized.However, they give a broad idea about the use of stance lexical bundles. Also, since the use of lexical bundles by nonnative speakers is not studied widely and there are very few studies conducted comparing the use of them by native and nonnative speakers focusing on (stance) lexical bundles, the results of this study could not be compared with other studies and generalizations could not be made.

5.4. Future Research

One of the advantages of corpus studies is being able to use data that is becoming larger and larger progressively. Therefore, since this study dealt with limited amount of data, in thefuture, by using a larger amount of data with more sophisticated and more easilyaccessible software tools, studies related to lexical bundles can be done more deeply providing more reliability, validity and objectivity. In order to do this, more corpora should be collected in various languages other than English.Nonnative data helps researchers to see how mother tongue and culture differences affect metadiscourse use. More nonnative corpora shouldbe collected to make accurate inferences about language learning and teaching by making better contrastive interlangauge analysis. These may help researchers understand human mind better at an abstract level. After understanding how these structures are processed and where the problematic areas are, more experimental studies can be conducted to help learners acquire these structures and become more fluent and proficient language learners. Moreover, in this study, only stance lexical bundles are analyzed. Studies looking at

other functions of lexical bundles should be conducted to understand these structures completely. Focusing other different genres, academic and nonacademic, spoken and written, would also give more insights on this topic.

REFERENCES

Abuhekama, G., Faraj, R., Feldman, A., Fitzpatrick, E. (2008). Annotating an Arabic Learner Corpus for Error. *Proceedings of International Conference on Language Resources and Evaluation,* 1347-1350. Marrakech, Morocco. Retrieved 6 December 2011 from: http://www.lrecconf.org/proceedings/lrec2008/pdf/343_paper.pdf

Ağçam, R. (2014). *A Corpus-Based study on author stance in academic English*. Unpublished doctoral thesis. Çukurova University, Adana.

Aijmer, K. (2009). *Corpora and language teaching.* Amsterdam, Philadelphia: John Benjamins

Allen, D. (2009). Lexical bundles in learner writing: An analysis of formulaic language in the ALESS Learner Corpus. *Komaba Journal of English Education*, 1, 105-127.

Anğ, F. (2006). *Effectiveness of corpus consultation through concordancing on the formulaic academic language use of freshman ELT students*. Unpublished Master's Thesis.Bogazici University, Istanbul.

Babanoğlu, P. (2012). *A Corpus-Based Study On Turkish EFLLearners'Written English: TheUse of Adverbial Connectors By Turkish Learners.* Unpublished Doctoral Thesis. Çukurova University, Adana.

Baker, P., Hardie, A., & McEnery, T. (2006). *A glossary of corpus linguistics*. Edinburg: Edinburg.

Bal, B (2010). *Analyses of Four-word lexical bundles in published research articles written by Turkish scholars.* Unpublished Master's Theses. Georgia State University, Georgia. Retrieved 5 july 2011 from: http://digitalarchive.gsu.edu/alesl_theses/2

Başal, A. (2006).*Use of first person pronouns: A corpus based study of journal articles*. Unpublished Masters Theses.Çukurova University.

Belz, J. &Vyatkina, V. (2008).The pedagogical mediation of a developmental learner corpus for classroom-based language instruction. *Language Learning and Technology*, 12 (3), 33-52.

Biber, D., Johansson, S., Leech, G., Conrad, S. &Finegan, E. (1999). *Longman grammar of spoken and written English*. London: Longman.

Biber, D., Conrad, S. & Cortes, V. (2004). If you look at: Lexical Bundles in University teaching and textbooks. *Applied Linguistics,* 25 (3), 371-405.

Biber, D. (2006). *University language: A corpus-based study of spoken and written registers.* Amstersam, Philadelphia: John Benjamins.

Biber, D., &Barbieri, F. (2007).Lexical Bundles in University spoken and written registers. *English for Specific Purposes, 26,* 263–286.

Chen, Y., & Baker, P. (2010).Lexical bundles in L1 and L2 academic writing. *Language Learning and Technology*, 14 (2), 30-49.

Chen, Z. (2012). Expression of epistemic stance in EFL Chinese University students' writing. *English Language Teaching*; 5 (10), 173-179.

Cheng, W., Warren, M. &Xun-Fen, X. (2003). The language learner as language researcher: Putting corpus linguistics on the timetable. *System,* 173-186.

Coates, J. (1990). Modal meaning: The semantics-pragmatic interference. *Journal of Semantics,* 7(1), 53-63.

Coates, J. (1995). The expression of root and epistemic possibility in English. In J. Bybee, & S. Fleischman (eds), *Modality in Grammar and Discourse*, 55-66. Amsterdam, Philadelphia: John Benjamins.

Cortes, V. (2002). Lexical bundles in Freshman Composition. In R. Reppen, S. M. Fitzmaurice and D. Biber (Eds), *Using Corpora to Explore Linguistic Variation*, 131-145: Amsterdam, Philadelphia: John Benjamins.

Cortes, V. (2004). Lexical bundles in published and student disciplinary writing: Examples from history and biology. *English for Specific Purposes, 23,* 397–423.

Cortes, V. (2006).Teaching lexical bundles in the disciplines: An example from a writing intensive history class. *Linguistics and Education*, 17, 391–406.

Cortes, V. (2008). A comparative analysis of lexical bundles in academic history writing in English and Spanish. *Corpora,* 3, 43-57.

Conrad, S. & Biber, B. (2004). The frequency and use of lexical bundles in conversation and academic prose.*LEXICOGRAPHICA,* 20, 56-71.

De Klerk, V. (2005). Procedural meanings of well in a corpus of Xhosa English. *Journal of Pragmatics,* 37(8), 1183-1205.

Ebeling, S.O. (2011). Recurrent Word-combinations in English student essay.*Nordic Journal of English Studies*, 10 (1), 46-76.

Flowerdew, L. (2001). The Exploitation of small learner corpora in EAP materials design. In Ghadessy, M., Henry, A. & Roseberry, R.L. (Eds.) *Small Corpus Studies and ELT Theory and Practice.* Amsterdam: John Benjamins.

Gajic, N. (2010). *Materials development: Corpus-based study of authentic texts for business.* Unpublished Master's Thesis. San Diego State University, California.

Girgin, U. (2011). *Corpus based activities at lower levels of EFL proficiency: The effectiveness of using concordance lines on grammar learning.* Unpublished Master's Thesis, Bilkent University, Ankara.

Granger, S. (1996) From CA to CIA and Back: An Integrated Approach to Computerized Bilingual and Learner Corpora. In K. Aijmer, B. Altenberg, & M. Johansson (Eds.) *Languages in Contrast.Text-based cross-linguistic studies* (pp. 37-51).Lund Studies in English 88. Lund: Lund University Press.

Granger, S. (1998). *Learner English on computer*. London & New York: Addison Wesley Longman.

Granger, S., Hung, J. &Petch-Tyson, S. (2002). *Computer learner corpora, second language acquisition and foreign language teaching*. Amsterdam, Philadelphia: John Benjamins.

Granger, S. (2002). *A Bird's-eye view of learner corpus research. in computer learner corpora, second language acquisition and foreign language teaching.* Granger, S. & Hung, J. &Petch-Tyson, S. (eds), 3-33. John Benjamins, Amsterdam.

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.

Granger, S., Dagneaux, E. & Meunier, F. (2009).*The international corpus of learner English version 2. handbook and CD-ROM*. Louvain-la-Neuve: PressesUniversitaires de Louvain.

Gries, S. (2009). What is corpus linguistics? *Language and Linguistics Compass*, 3, 1–17.

Guo, X. (2006).*Verbs in the Written English of Chinese Learners: A Corpus-based Comparison between Non-native Speakers and Native Speakers*. Unpublished Doctors Theses.University of Birmingham.

He, A. (2008).Application of linguistics to EFL teacher education in China. *Proceedings of TACL 8th Conference.* 339-345. Lisbon

Herbel-Eisemann, B., Wagner, D. & Cortes, W. (2010). Lexical bundle analysis in mathematics classroom discourse: The significance of stance. *Educ Stud Math*, 75, 23-42.

Hernandez. P.S. (2013). Lexical bundles in three oral corpora of university students.*Nordic Journal of English Studies*, 13(1), 187-209.

Holmes, J. (1982). Expressing doubt and certainty in English. *RELC Journal*, 13(2), 9-28.

Hua, C. & Quifang, W. & Aijun, L. (2008).A Learner Corpus-ESCCL. Proceedings of the Speech Prosody Conference (edited Barbosa, P.A., Madureira, S. & Reis, C). Campinas, Brazil. Retrieved 8 September 2013 from: http://sprosig.isle.illinois.edu/sp2008/papers/id187.pdf

Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge: Cambridge University Press.

Hyland, K. (2002). Authority and invisibility: Authorial identity in academic writing. *Journal of Pragmatics,* 34, 1091–1112.

Hyland, K. (2005). Stance and engagement: A model of interaction in academic dicourse. *Discourse Studies*, 7, 2, 173-192.

Hyland, K. (2008, a). As can be seen: Lexical bundles and disciplinary variation. *English for Specific Purposes,* 27, 4–21

Hyland, K. (2008, b). Academic clusters: Text patterning in published and postgraduate writing. *International Journal of Applied Linguistics*, 18 (1), 41-62.

Ishikawa, S. (2009). Phraseology overused and underused by japanese learners of English: A contrastive interlanguage analysis phraseology. *Corpus Linguistics and Lexicography*, 87-100.

Jablonkai R.R. (2010). *A Corpus-linguistic investigation into the lexis of written English EU Discourse: An ESP Pedagogic Perspective.* Unpublished doctoral thesis.EötvösLorándUniversity, Budapest.

Jafarpour, A.K., Hashemian, M. &Alipour, S. (2013). A corpus based approach toward teaching collocation of synonyms. *Theory and Practice in Language Studies*, 3 (1), 51-60.

Jalali, H., Rasekh, A.E., &Rizi, M.T. (2008).Lexical bundles and intradisciplinary variation: The case of applied linguistics. *International Journal of Language Studies*, 2 (4), 447-484.

Juknevičienė, R. (2011). *Lexical bundles in non-native speaker and native speaker written English.* Unpublished doctoral dissertation. Vilnius University.

Kamimura, T. & Oi, K. (1998). Argumentative strateies in American and Japanese English. *World Englishes*, 17 (3), 307-323.

Kim, Y.J. (2009). Korean lexical bundles in conversation and academic texts. *Corpora*, 4(2), 135-165.

Krieger, D. (2003). Corpus linguistics: What is it and how it can be applied to teaching. *The Internet TESL Journal*, 9 (3). Retrieved 9 March 2014, from: http://iteslj.org/Articles/Krieger-Corpus.html

Liu, D. & Jiang, P. (2009).Using a Corpus-Based lexicogrammatical approach to grammar instruction in EFL and ESL contexts.*Modern Language Journal*, 93 (i), 61-78.

Louwerse, M.M. & McCarthy, P.M. McNamara, D.S, & Graesser, A.C. (2004).Variation in Language and Cohesion across Written and Spoken Registers. In K. Forbus, D. Gentner& T. Regier (Eds.), *Proceedings of the 26th Annual Meeting of the Cognitive Science Society* (pp. 843–848). Mahwah, NJ: Erlbaum.

McEnery, T. & Wilson, A. (2001).*Corpus Linguistics (Second Ed.).*Edinburg University Press.

McEnery, A., Baker, P. & Wilson, A. (1995).A Statistical Analyses of Corpus Based vs. Traditional Human Teaching Methods of Parts of Speech Analysis.*Computer Assisted Language Learning*, 8 (2/3), 259-274 (cited in McEnery and Wilson, 2001).

Mukherjee, J. (2006). Corpus linguistics and language pedagogy: The state of the art – and beyond. In S. Braun, K. Kohn & J. Mukherjee (Eds.), *Corpus Technology and Language Pedagogy: New Resources*, New Tools, New Methods (5-24). Frankfurt: Peter Lang.

Neely E. & Cortes.V. (2009). A little bit about: Analyzing and Teaching Lexical Bundles in Academic Lectures. *Language Value*, 1(1), 17-38. Servei de Publicacions de la UniversitatJaume I: Castelló, Spain. Retrieved 9 July 2012, from: http://www.erevistesuji.es/languagevalue.

Nekrasova, T.M. (2009). English L1 and L2 Speakers' Knowledge of Lexical Bundles. *Language Learning*, 59 (3), 647-686.

Newman, J. & Rice, S. (2004). Patterns of usage for English SIT, STAND, and LIE: A Cognitively Inspired Exploration in Corpus Linguistics. *Cognitive Linguistics*, 15(3), 351–396.

Nesi, H., &Basturkmen, H. (2006). Lexical Bundles and Discourse Signaling in Academic Lectures.*International Journal of Corpus Linguistics, 11,* 283-304.

Ohta, A.S. (1991). Evidentiality and Politeness in Japanese. *Issues in Applied Linguistics*, 2(2), 211-238.

Rayson, P., & Garside, R. (2000).Comparing Corpora using Frequency Profiling.*Proceedings of the Workshop on Comparing Corpora*, 9, 1-6. Stroudsburg, PA, USA.

Reder, S., Harris, K &Setzler, K. (2003).The Multimedia Adult learner Corpus.*TESOL Quarterly*, 37 (3), 546-557.

Rodriguez, I.R. & Moreno, E.M.G. (2009). Teaching Idiomatic Expressions to Learners of EFL Through a Corpus Based on Disney Movies. A survey of corpus-based research [Recursoelectrónico].240-253. Retrieved 8 october 2014, from: http://www.um.es/lacell/aelinco/contenido/pdf/17.pdf

Sahillioğlu, F., Şahinkayası, Y. &Şahinkayası, H. (2011). *Effectiveness of Two Corpus-Based Approaches on Turkish Students' Learning English Grammar.5th International Computer and Instructional Technologies Symposium.* FıratUniversity, Elazığ. Retrieved 8 October 2014, from: http://web.firat.edu.tr/icits2011/papers/27712.pdf

Saeed, N.A. &Waly, S. (2009). Corpus for Classrooms: Ideas for Material Design.Proceedings of the 10th METU ELT Convention. METU, Ankara. Retrieved 9 October 2014, from: http://dbe.metu.edu.tr/convention/proceedingsweb/Corpus.pdf

Scott, M. (2001). Comparing Corpora and Identifying Key Words, Collocations, and Frequency Distributions Through the WordSmith Tools Suite of Computer Programs. In M. Ghadessy, A. Henry & R. Roseberry (Eds.) *Small Corpus Studies and ELT* (pp. 47-67). Amsterdam: Benjamins.

Scott, M. (2008).WordSmith Tools version 5, Liverpool: Lexical Analysis Software.

Scott, M. (2010). *Introduction to WordSmith Tools.* Retrieved 20 December 2014, from: http://www.lexically.net/downloads/version5/HTML/?getting_started.htm

Shih, R.H. (2000).Compiling Taiwanese Learner Corpus of English.*Computational Linguistics and Chinese Language Processing*, 5(2), 87-100.

Sinclair, J. (2001). Preface. In Ghadessy, M., Henry, A. & Roseberry, R.L. (Eds.) *Small Corpus Studies and ELT : Theory and Practice.* Amsterdam: John Benjamins.

Strunkyte, G. &Jurkūnaite, E. (2008).*Written academic discourse: lexical bundles in humanities and natural sciences*. Unpublished BA thesis. Retrieved 9 July 2012, from: http://www.lexically.net/wordsmith/corpus_linguistics_links/Lexical%20Bundles%20i_n%20Humanities%20and%20Natural%20Sciences.pdf

Smith C., Butler, N.L., Griffith, K.G. &Kritsonis, W.A. (2007).The Role of Communication Context, Corpus-based Grammar, and Scaffolded Interaction in ESL/EFL Instruction.*The Lamar UniversityElectronic Journal of Student Research.* Retrieved 5 September 2014, from: http://files.eric.ed.gov/fulltext/ED495290.pdf

Şanal, F. (2007).*A learner corpus based study on second language lexicology of Turkish students of English*. Unpublished Doctors Theses, Çukurova University, Adana.

Tagnin, Stella E. O. (2006). *language and computers, corpus linguistics around the World*. Edited by Andrew Wilson, Dawn Archer, Paul Rayson, pp. 195-202(8).

Taki, S. &Jafarpour, F.(2012). Engagement and stance in academic writing: A study of English and persian research articles. *Mediterranean Journal of Social Sciences*, 3 (1), 157-168.

Tenuta, A. M., Oliveira, A. L. A. M. &Orfanó, B. M. (2012). How brazilian learners express modality in their writing: A corpus-based study on lexical bundles. *RevistaIntercâmbio*, 5.(26), 1-15, 2012. São Paulo: LAEL/PUCSP.

Thomas, D. (2005). Type Token Ratios in One Teacher's Classroom Talk: An Investigation of Lexical Complexity. Retrieved 16 December 2014, from: http://www.birmingham.ac.uk/Documents/collegeartslaw/cels/essays/languageteaching/DaxThomas2005a.pdf

Tono, Y., Kaneko, T., Isahara, H., Izumi, E., Saiga, T. & Kaneko, E. (2001).Developing a One Million Word Spoken EFL Learner Corpus.*PAC3 at JALT Conference Proceedings*, 872-880.Kitakyushu, Japan.

Uysal, H. H. (2012). Argumentation across L1 and L2 Writing: Exploring CulturalInfluences and Transfer Issues. *Vigo International Journal of Applied Linguistics*, 9, 133-159.

Wei, Y. & Lei, L. (2011). Lexical bundles in the academic writing of advanced chinese EFL learners. *RELC Journal*, 42 (2), 155-166.

Williamson, G. (2013). http://www.sltinfo.com/type-token-ratio.html Retrieved 10 July 2012, from:

Williamson, G. (2009). http://www.sltinfo.com/wp-content/uploads/2014/01/type-token-ratio.pdf Retrieved 3 August 2013, from:

Wu, W.S. (?) The Integration of Corpus-Based Data into Grammar Instruction: Using Advise, Recommend, and Suggest as an Example. Retrieved 5 October 2014, from: http://web.chu.edu.tw/~wswu/publications/papers/journals/07.pdf

APPENDICES

Appendix A. LOCNESS Results

Stance Lexical Bundles	Frequency
THE FACT THAT	71
BE ABLE TO	50
HAVE TO BE	23
ARE/IS GOING TO	22
I FEEL THAT	20
DO NOT WANT	18
BE ALLOWED TO	15
NEED TO BE	15
WILL CONTINUE TO	14
THE OPPONENTS OF	13
I BELIEVE THAT	12
THERE WOULD BE	12
WOULD HAVE TO	12
IT IS IMPORTANT	11
WOULD LIKE TO	11
IS/ARE FORCED TO	10
WOULD HAVE BEEN	10
ARGUE THAT THE	9
DO NOT HAVE TO	9
I KNOW I/THE/YOU	9
I THINK+N (IT/THE)	9
I THINK THAT	9
IN MY OPINION	9
IT IS OBVIOUS	9
THERE WILL BE	9
CAN BE SEEN	8
I HAVE SEEN	8
PEOPLE FEEL THAT	8
SHOULD BE ALLOWED TO	8
WILL BE ABLE TO	8
WILL HAVE TO	8
WOULD BE ABLE TO	8
ARE WILLING TO	7
CONSIDERED TO BE	7

I WOULD LIKE TO	7
IS/ARE ALLOWED TO	7
IS VERY IMPORTANT	7
IT IS POSSIBLE	7
MAY HAVE BEEN	7
ONE CAN SEE	7
SHOULD BE ABLE TO	7
THE BEST WAY TO	7
THE CASE AGAINST THE	7
THE RESULTS OF	7
THIS ARGUMENT IS	7
THIS CAN BE	7
WE NEED TO	7
ARE EXPECTED TO	6
DO NOT REALIZE	6
I KNOW THAT	6
IS A PROBLEM	6
IS AN ISSUE	6
IS SUPPOSED TO	6
IS THE BEST	6
IT COULD BE	6
IT IS CLEAR THAT	6
IT IS HARD	6
PEOPLE NEED TO	6
REALIZE THAT THE/THIS	6
SHOULD NOT BE ALLOWED	6
THE OPPONENTS TO	6
THERE SHOULD BE	6
WE HAVE SEEN	6
WE HAVE TO	6
A GOOD EXAMPLE OF	5
APPEAR TO BE	5
ARGUMENT AGAINST THE	5
ARGUMENT IN FAVOR OF	5
DOES NOT MEAN THAT	5
I FEEL THE	5
IS NOT TRUE	5
IT IS IMPORTANT FOR	5
IT IS IMPORTANT TO	5
IT IS TRUE	5

IT SEEMS THAT	5
MUST DEAL WITH	5
PEOPLE SEE THE	5
SHOULD HAVE BEEN	5
THE FACT IS	5
THIS COULD BE	5
THIS WOULD BE	5
WILL ALWAYS BE	5
ARE SUPPOSE TO	4
CAN ALSO BE	4
COULD ALSO BE	4
COULD HAVE BEEN	4
DO NOT KNOW	4
DOES NOT SEEM	4
HAVE SHOWN THAT	4
IS WRONG AND	4
IT IS COMMON	4
IT IS DIFFICULT TO	4
IT IS NOT FAIR	4
IT IS TRUE THAT	4
IT MAY BE	4
IT WILL BE	4
IT WOULD BE A	4
LESS LIKELY TO	4
MORE LIKELY TO	4
PEOPLE THINK THAT	4
PEOPLE WANT TO	4
PEOPLE WOULD BE	4
SAY THAT THE	4
SHOULD HAVE A	4
SOME PEOPLE FEEL	4
SOME PEOPLE WOULD	4
THE BELIEF THAT	4
THE BEST SOLUTION	4
THERE MUST BE	4
THERE WILL BE NO	4
THIS IS NOT TRUE	4
THIS MEANS THAT	4
THIS SHOWS THAT	4
TO ME THAT	4

WE KNOW IT	4
WOULD NOT HAVE BEEN	4
ADVOCATES OF THE	3
AS WE KNOW	3
BY THE FACT THAT	3
CAN BE EASILY	3
CAN BE SHOWN	3
CAN SAY THAT	3
DO NOT BELIEVE IN	3
DO NOT UNDERSTAND	3
DOES THIS MEAN THAT	3
HAVE TO AGREE	3
I AGREE THAT	3
I AGREE WITH	3
I AM SURE	3
I DON'T BELIEVE	3
I HAVE HEARD	3
I UNDERSTAND THAT	3
I'M ABLE TO	3
I'M GOING TO	3
IS A BAD	3
IS A COMMON	3
IS A SERIOUS	3
IS AGAINST THE	3
IS EASY TO	3
IS NOT NECESSARY	3
IS SEEN AS	3
IT HAS BEEN PROVEN	3
IT IS EASY	3
IT IS POSSIBLE THAT	3
IT SEEMS AS	3
IT SEEMS TO ME	3
IT WOULD SEEM	3
IT'S IMPORTANT THAT THE	3
MANY PEOPLE BELIEVE	3
MUST BE TAKEN	3
MUST DEAL WITH THE	3
MUST LOOK AT	3
MUST REALIZE THAT	3
NEED TO HAVE A	3

NOT ALLOWED TO	3
NOT FAIR THAT	3
ON THE FACT THAT	3
PEOPLE ARGUE THAT	3
PEOPLE BELIEVE THAT	3
PEOPLE DO NOT REALIZE	3
PEOPLE SAY THAT	3
PERHAPS IT IS BECAUSE	3
PROBABLY WOULD NOT	3
SEE IT AS A	3
SEEMED TO BE	3
SHOULD BE CONSIDERED	3
SHOULD NOT HAVE TO	3
SIGNIFICANT IMPACT ON	3
THAT IS CONSIDERED	3
THAT MAY BE	3
THE CLAIM IS	3
THE CONTROVERSY IS	3
THE FACT IS THAT	3
THE HOPE THAT	3
THERE CAN BE	3
THINK TWICE BEFORE	3
THIS IS IMPORTANT	3
THIS IS TRUE	3
THIS WAS UNFAIR	3
WE ALL KNOW	3
WE REALIZE THAT	3
WITHOUT A DOUBT	3
YOU DON'T HAVE TO	3
YOU HAVE TO	3

Appendix B. TIICLE Results

Stance Lexical Bundles	Frequency
I THINK+N (THE/IT/THIS,etc)	88
IN MY OPINION	70
I WANT TO	59
BE ABLE TO	53
I BELIEVE THAT	43
IT CAN BE	39
DO NOT WANT	36
IT WILL BE	32
IS/ARE VERY IMPORTANT	31
THE FACT THAT	31
THERE WILL BE	31
WE CAN SEE	30
I THINK THAT	29
DO NOT KNOW	27
DO NOT THINK	30
THERE SHOULD BE	24
WE CAN SAY	23
WE SEE THAT	23
WE HAVE TO	22
IS THE MOST IMPORTANT	21
IS/ARE GOING TO	19
IT MEANS THAT	19
DO NOT HAVE TO	18
CAN BE DONE	17
I CAN SAY	17
IS/ARE NOT IMPORTANT	17
THERE IS NO NEED	17
PEOPLE THINK THAT	16
WE KNOW THAT	16
IN FACT THERE/THIS/THE	15
IT IS NECESSARY	15
SEEMS TO BE	15
THE IMPORTANT THING IS	15
WE CAN SAY THAT	15
ACCORDING TO ME	14
DO NOT UNDERSTAND	13

IT IS BETTER	13
IT IS TRUE	13
SEEN AS A/AN	13
THIS MEANS THAT	13
AGREE WITH THE\THIS	12
CAN BE USED	12
I CAN SAY THAT	12
I HAVE TO	12
I WILL BE	12
IN FACT IT IS	12
IS/ARE SUPPOSED TO	12
IT IS POSSIBLE	12
CAN SEE THAT	11
DO NOT NEED TO	11
IT IS NOT ENOUGH	11
THERE MUST BE	11
WILL HAVE TO	11
WOULD LIKE TO	11
DO NOT BELIEVE	10
DO NOT FEEL	10
I CAN NOT	10
I DO NOT KNOW	10
IT IS A FACT	10
IT IS OBVIOUS THAT	10
IT IS SAID THAT	10
IT IS TRUE THAT	10
WANT TO SAY	10
WE THINK THAT	10
ARE THE MOST IMPORTANT	9
CAN BE SEEN	9
IT IS VERY DIFFICULT	9
IS/ARE BETTER THAN	9
NEED TO BE	9
SOME PEOPLE THINK	9
WE NEED TO	9
WE SEE THE	9
WE WANT TO	9
ARE ABLE TO	8
DO NOT SHOW	8
I DO NOT AGREE	8

I DO NOT THINK	8
IS NOT EASY	8
IT IS CLEAR THAT	8
IT IS SEEN	8
ARE MORE IMPORTANT THAN	7
ARE NOT ABLE TO	7
CAN NOT SEE	7
CONSIDERED TO BE	7
DOES NOT MEAN	7
IMPORTANT POINT IS	7
IMPORTANT THING IS THAT	7
IMPORTANT THING IS TO	7
IS THE REASON	7
IT IS ALWAYS	7
IT IS IMPOSSIBLE TO	7
IT IS NOT IMPORTANT	7
IT MAY BE	7
IT MUST BE	7
MIGHT NOT BE	7
PEOPLE BELIEVE THAT	7
PEOPLE SAY THAT	7
SHOULD BE CONSIDERED	7
WE CAN NOT SAY	7
TELL US THAT	7
THERE WILL BE NO	7
ARE SEEN AS	6
AS IT IS KNOWN	6
AS WE ALL KNOW	6
AS WE SEE	6
CAN BE FOUND	6
CAN T THINK	6
CONSIDERED AS A	6
COULD NOT BE	6
DO NOT LIKE	6
I AM GOING TO	6
I WANT TO SAY	6
I WANT TO TALK	6
I WOULD LIKE TO	6
IS THE ONLY WAY	6
IS THE RESULT OF	6

IS THE WORST	6
IT IS BETTER TO	6
IT IS CERTAIN THAT	6
IT IS HARD TO	6
IT WOULD BE	6
SEEM TO BE	6
THERE WOULD BE NO	6
WE CAN SEE THAT	6
WE SHOULD CONSIDER	6
WE THINK ABOUT	6
DO NOT ACCEPT	5
I CAN ACCEPT	5
I KNOW THAT	5
I MEAN THE	5
I WANT TO EXPLAIN	5
IS A BIG PROBLEM	5
IT IS HARMFUL	5
IT IS NOT POSSIBLE	5
IT IS NOT RIGHT	5
IT IS POSSIBLE TO	5
IT IS VERY IMPORTANT	5
IT IS WRONG	5
IT SEEMS THAT	5
SHOULD BE ABLE TO	5
SHOULD BE BANNED	5
THE BEST SOLUTION	5
THE MAIN PROBLEM IS	5
THE ONLY SOLUTION	5
THE POSSIBILITY OF	5
THERE MAY BE	5
THIS IS A FACT	5
THIS SHOWS THAT	5
WE ALL KNOW THAT	5
WE WILL NOT	5
WILL ALWAYS BE	5
WILL BE ABLE TO	5
WILL SEE THAT	5
A GOOD EXAMPLE	4
ARE AGAINST THE	4
ARE ALLOWED TO	4

ARE MORE POWERFUL THAN	4
ARE MORE SUPERIOR THAN	4
ARE NOT ALLOWED TO	4
ARE STRONGER THAN	4
ARE THE BEST	4
ARE USEFUL FOR	4
AS EVERYBODY KNOWS	4
CAN ALSO BE	4
CAN BE ACCEPTED	4
CAN BE DISCUSSED	4
CAN NOT DECIDE	4
DO NOT LET	4
DO NOT SAY	4
DOES NOT MEAN THAT	4
EVERYONE IS SUPPOSED TO	4
HAVE TO OBEY THE	4
I AM SURE THAT	4
I CLAIM THAT	4
I DO NOT MEAN	4
I MEAN THAT	4
I WILL TRY TO	4
IS INDISPENSABLE FOR	4
IS NOT SO EASY	4
IS NOT SO IMPORTANT	4
IS NOT TRUE	4
IT IS CLAIMED THAT	4
IT IS DIFFICULT TO	4
IT IS EASY TO	4
IT IS IMPORTANT TO	4
IT IS MEANINGLESS	4
IT IS NOT EASY	4
IT IS UNNECESSARY	4
IT MAY HAVE	4
IT MIGHT BE	4
MANY PEOPLE THINK	4
PEOPLE HAVE TO	4
PEOPLE SHOULD THINK	4
SHOULD BE ALLOWED	4
SHOULD BE MENTIONED	4
SHOULD BE SAID THAT	4

SOME PEOPLE BELIEVE	4
THERE ARE CERTAIN	4
THERE IS AN IMPORTANT	4
THERE IS NO REASON	4
THERE IS NO WAY	4
UNDERSTAND THE IMPORTANCE OF	4
WE CAN EASILY SEE	4
WE CAN UNDERSTAND THAT	4
WE CONSIDER THE	4
WE SHOULD THINK	4
WE THINK OF	4
A GOOD WAY TO	3
ALWAYS CLAIM THAT	3
APPARENTLY SOME PEOPLE JUST	3
ARE AGAINST TO	3
ARE LIKELY TO	3
ARE OBLIGED TO	3
ARE REGARDED AS	3
CAN BE SAID	3
CAN CLAIM THAT	3
DO WE THINK THAT	3
I ACCEPT THE	3
I ADMIT THAT	3
I AGREE WITH	3
I ALSO BELIEVE THAT	3
I ARGUE THAT	3
I CAN NOT SAY	3
I DO NOT WANT	3
I HAVE SEEN	3
I KNOW IT	3
I SUPPORT THE	3
IS ALWAYS SUPERIOR TO	3
IS CONSIDERED TO	3
IS EASILY SEEN	3
IS NOT ACCEPTED	3
IS NOT DIFFICULT	3
IS NOT SEEN	3
IS NOT SUFFICIENT FOR	3
IS NOT SUITABLE FOR	3
IS SO CLEAR THAT	3

IS THE MOST VALUABLE	3
IS USELESS AND	3
IT CAN BE SEEN	3
IT DOESN T MEAN	3
IT IS A BAD	3
IT IS ACCEPTABLE	3
IT IS ACCEPTED THAT	3
IT IS ARGUED THAT	3
IT IS BELIEVED THAT	3
IT IS BETTER FOR	3
IT IS DIFFICULT FOR	3
IT IS EASILY	3
IT IS FORBIDDEN	3
IT IS ILLEGAL	3
IT IS IMPOSSIBLE FOR	3
IT IS INEVITABLE TO	3
IT IS KNOWN THAT	3
IT IS NONSENSE TO	3
IT IS POSSIBLE FOR	3
IT IS SEEN THAT	3
IT IS STATED THAT	3
IT IS THE BEST	3
IT IS THOUGHT THAT	3
IT IS VERY EASY	3
IT MUST NOT BE	3
IT S CONSIDERED THAT	3
IT S GOOD FOR	3
IT SEEMS TO	3
MUST BE CONSIDERED	3
NOBODY CAN BE	3
NOBODY CAN PREVENT	3
PEOPLE MAY THINK THAT	3
SHOULD NOT BE FORGOTTEN	3
SHOULD THINK ABOUT	3
SOME PEOPLE BELIEVE THAT	3
SOME PEOPLE SAY	3
SOME SAY THAT	3
SUPPORT THE IDEA THAT	3
THERE IS NO DOUBT	3
THIS SEEMS TO BE	3

THIS SHOWS THE	3
WANT TO KNOW	3
WE CAN TALK ABOUT	3
WE DO NOT KNOW	3
WE MUST KNOW	3
WE SHOULD KNOW THAT	3
WE SHOULD NOT FORGET	3
WE SHOULD TRY TO	3
WILL ALSO BE	3
WILL NEED TO	3
WILL SHOW THAT	3
WOULD BE BETTER	3

Appendix C. JPICLE Results

Stance Lexical Bundles	Frequency
I THINK+N (IT/THE/THIS,etc)	309
I THINK THAT	279
I WANT TO	185
WE HAVE TO	108
WE NEED TO	94
IS/ARE VERY IMPORTANT	68
I DO NOT THINK	67
IT IS IMPORTANT	48
DON'T HAVE TO	43
WE WANT TO	40
I BELIEVE THAT	38
I WOULD LIKE TO	38
IT WILL BE	38
I DON'T WANT	34
ARE ABLE TO	32
IT IS GOOD	32
MOST IMPORTANT THING	31
WILL BE ABLE TO	31
HAVE TO BE	30
IS VERY DIFFICULT	30
IT IS DIFFICULT	30
WANT TO SAY	30
I DON'T KNOW	28
IS VERY USEFUL	28
IT IS SAID	28
I NEED TO	25
IT IS NATURAL	25
IT IS SAID THAT	25
IS GOOD FOR	24
IS THE BEST	23
IT IS NECESSARY	23
IT MAY BE	23
THE FACT THAT	23
IT IS VERY IMPORTANT	22
I AM GOING TO	21
IT IS TRUE	21

IT MIGHT BE	21
WILL NOT BE	21
DON'T NEED TO	20
IN MY OPINION	20
IS THE MOST IMPORTANT	20
I HAVE TO	19
IS NOT GOOD	19
IT IS IMPORTANT TO	19
IMPORTANT THING IS	18
IT IS DIFFICULT TO	17
IT SEEMS THAT	17
I DON'T THINK SO	16
IS NECESSARY FOR	16
IT IS TRUE THAT	16
SOME PEOPLE SAY	16
I DON'T THINK THAT	15
IT IS EASY TO	15
IT IS IMPORTANT FOR	15
IT IS VERY DIFFICULT	15
THE BEST WAY	15
ARE GOING TO	14
I CAN NOT	14
IS/ARE NOT ENOUGH	14
MAY NOT BE	14
TEND TO THINK	14
TO ME THAT	14
WE DON'T KNOW	14
ARE VERY USEFUL	13
I HAVE HEARD	13
I HEARD THAT	13
I THINK SO	13
IS MORE IMPORTANT	13
IT IS GOOD TO	13
IT IS VERY USEFUL	13
IT MEANS THAT	13
PEOPLE NEED TO	13
I KNOW THE/I	12
IT CAN BE	12
WE CAN ALSO	12
WE CAN UNDERSTAND	12

I BELIEVE IT/THE/THERE	11
IF WE WANT TO	11
IT IS BETTER TO	11
IT IS CLEAR	11
IT IS DIFFICULT FOR	11
IT IS GOOD FOR	11
IT IS NECESSARY TO	11
IT IS POSSIBLE	11
SEEM TO BE	11
SOME PEOPLE SAY THAT	11
WE DON'T HAVE TO	11
I AM SURE THAT	10
I HOPE THAT	10
IS MOST IMPORTANT	10
IT SEEMS TO ME	10
KNOW HOW TO	10
MOST IMPORTANT THING IS	10
WE ARE ABLE TO	10
I CAN SAY	9
I DON'T AGREE	9
I HOPE/I/THE/WE	9
I SUPPOSE THAT	9
IS VERY DIFFICULT FOR	9
IT IS IMPOSSIBLE TO	9
MAY THINK THAT	9
THERE WILL BE	9
I AM ABLE TO	8
I HAVE HEARD THAT	8
I WILL BE	8
I WOULD BE	8
IS A GREAT	8
IT IS CLEAR THAT	8
IT IS HARD	8
IT IS NATURAL THAT	8
IT IS NOT GOOD	8
IT IS WRONG	8
IT WOULD BE	8
WE CANNOT SAY	8
I AGREE WITH	7
I AM SURE	7

IS A PROBLEM	7
IS USEFUL FOR	7
LOOK AT THE	7
MOST IMPORTANT THING FOR	7
THE ONLY WAY	7
WE CAN SAY THAT	7
WE DON'T NEED TO	7
WE HAD BETTER	7
WE THINK THAT	7
ARE SUPPOSED TO	6
ARE TOO EASY	6
I CAN'T UNDERSTAND	6
I KNOW THAT	6
IN FACT I	6
IS A COMMON	6
IS A GOOD WAY	6
IS ALSO IMPORTANT	6
IS EASY FOR	6
IS USEFUL TO	6
IS VERY BAD	6
IS VERY DANGEROUS	6
IS VERY HARD	6
IS VERY INTERESTING	6
IT CAN BE SAID	6
IT IS CERTAIN THAT	6
IT IS IMPORTANT THAT	6
IT IS NATURAL FOR	6
IT IS NOT EASY	6
IT IS NECESSARY FOR	6
IT IS POSSIBLE TO	6
IT MUST BE	6
MAY BE ABLE TO	6
MAY SAY THAT	6
SHOULD BE ABLE TO	6
THE POSSIBILITY OF	6
WE MUST NOT FORGET	6
WE THINK ABOUT	6
ARE LIKELY TO	5
ARE NECESSARY FOR	5
I AM AGAINST	5

I CAN UNDERSTAND	5
I DISAGREE WITH	5
I DON'T KNOW WHY	5
I DON'T MEAN THAT	5
I SUPPORT THE	5
IS NOT EASY	5
IS REGARDED AS	5
IS VERY CONVENIENT	5
IS VERY DIFFICULT TO	5
IT IS HARD TO	5
IT IS USEFUL	5
MANY PEOPLE THINK	5
MAYBE IT IS	5
MIGHT NOT BE	5
MIGHT THINK THAT	5
PEOPLE MAY FEEL	5
PEOPLE THINK THAT	5
SHOULD NOT BE ALLOWED	5
SOME PEOPLE THINK	5
THE ONLY WAY TO	5
THERE MUST BE	5
THIS IS THE REASON	5
THIS IS TRUE	5
THIS IS VERY DIFFICULT	5
THIS MAY BE	5
THIS MEANS THAT	5
WANT TO THINK ABOUT	5
WE MUST THINK	5
WE SHOULD KNOW	5
WOULD LIKE TO SAY	5
A GOOD WAY TO	4
A SERIOUS PROBLEM	4
ARE EAGER TO	4
ARE NOT ABLE TO	4
AS WE KNOW	4
CAN BE SEEN	4
DO NOT FEEL	4
I ALSO THINK THAT	4
I CAN SAY THAT	4
I CAN'T SAY	4

I CONSIDER THAT	4
I DON'T AGREE WITH	4
I DON'T NEED TO	4
I DON'T SAY	4
I WILL THINK	4
I WISH THAT	4
I'D LIKE TO WRITE	4
IS A FACT THAT	4
IS A STRONG	4
IS A VERY IMPORTANT	4
IS BETTER THAN	4
IS NO NEED TO	4
IS NOT SO GOOD	4
IS NOT USEFUL	4
IS THE WORST	4
IT CANNOT BE	4
IT COULD BE	4
IT IS A FACT	4
IT IS AN IMPORTANT	4
IT IS COMMON	4
IT IS CONVENIENT	4
IT IS EASIER	4
IT IS MORE IMPORTANT	4
IT IS MOST IMPORTANT	4
IT IS NATURAL TO	4
IT IS SO HARD	4
IT IS SURE THAT	4
IT IS WONDERFUL	4
IT MAY BE DIFFICULT	4
MOST IMPORTANT THING IN	4
NEED TO CONSIDER	4
NEED TO THINK	4
NEED TO UNDERSTAND	4
PEOPLE TEND TO THINK	4
SHOULD BE BANNED	4
SHOULD NOT FORGET	4
SHOULD THINK ABOUT	4
THE FACT IS THAT	4
THE IMPORTANT THING IS	4
THERE IS A POSSIBILITY	4

THERE SHOULD BE	4
THIS CAN BE	4
WE CAN'T UNDERSTAND	4
WE DON'T THINK	4
WE DON'T UNDERSTAND	4
WE DON'T WANT TO	4
WE HAVE TO THINK	4
WE MAY NOT	4
WILL FEEL THE	4
YOU HAVE TO DO	4
A GOOD EXAMPLE	3
A GOOD IDEA	3
ALSO SAY THAT	3
ARE APT TO	3
ARE MORE IMPORTANT	3
ARE THE MOST IMPORTANT	3
CAN ALSO BE	3
CERTAINLY IT IS	3
DO NOT KNOW ABOUT	3
DO WE HAVE TO	3
I ADMIT THAT	3
I ALSO WANT TO	3
I AM NOT ABLE	3
I AM OPPOSED TO	3
I ASSUME THAT	3
I CAN KNOW	3
I CAN SEE	3
I CAN TELL	3
I CAN'T UNDERSTAND WHY	3
I DON'T AGREE THAT	3
I DON'T BELIEVE	3
I DON'T UNDERSTAND	3
I GUESS THAT	3
I INSIST THAT	3
I REALLY THINK	3
I REALLY WANT	3
I STRONGLY DISAGREE WITH	3
I THINK ABOUT	3
I THINK BECAUSE	3
I THINK IT'S BETTER	3

I WANT TO KNOW	3
I WANT TO SEE	3
I WILL THINK ABOUT	3
I WISH I	3
I'M NOT SURE	3
IMPORTANT THING IS THAT	3
IN MY VIEW	3
IS A SERIOUS	3
IS ALSO TRUE	3
IS BETTER FOR	3
IS CONVENIENT FOR	3
IS EASIER THAN	3
IS EASIER TO	3
IS HARD FOR	3
IS NO DOUBT	3
IS NOT ALWAYS TRUE	3
IS NOT CORRECT	3
IS NOT SIMPLE	3
IS NOT SO DIFFICULT	3
IS NOT SUITABLE FOR	3
IS NOT SURE	3
IS NOT TRUE	3
IS RATHER EASY	3
IT CAN SAY THAT	3
IT DOESN'T MEAN THAT	3
IT IS ALSO SAID	3
IT IS BORING	3
IT IS EASY FOR	3
IT IS ENOUGH	3
IT IS ESSENTIAL FOR	3
IT IS ILLEGAL	3
IT IS INDISPENSABLE TO	3
IT IS INTERESTING	3
IT IS MORE DIFFICULT	3
IT IS NOT STRANGE	3
IT IS SO DIFFICULT	3
IT IS SO IMPORTANT	3
IT IS STRANGE	3
IT IS USEFUL TO	3
IT IS USELESS	3

IT IS VERY INTERESTING	3
IT IS VERY NICE	3
IT IS WRONG TO	3
IT SHOWS THAT	3
IT WILL BE EASY	3
IT'S A REALLY IMPORTANT	3
IT'S IMPORTANT FOR	3
IT'S MORE IMPORTANT TO	3
IT'S NEEDED TO	3
MAY TEND TO THINK	3
MAYBE I THINK	3
MIGHT BE ABLE TO	3
MUST BE ABLE TO	3
MUST BE THINKING	3
MUST THINK ABOUT	3
MY OPINION IS THAT	3
NEVER BE ABLE TO	3
PEOPLE DON'T HAVE TO	3
PEOPLE DON'T KNOW	3
PEOPLE DON'T WANT TO	3
PEOPLE MAY SAY THAT	3
PEOPLE MUST BE	3
SHOULD BE DECIDED	3
SHOULD BE DONE	3
SHOULD KNOW ABOUT	3
SOMEONE SAY THAT	3
SUPPORT THIS IDEA	3
THERE IS A PROBLEM	3
THERE IS NO NEED	3
THERE IS POSSIBILITY	3
WE CAN SEE MANY	3
WE KNOW THAT	3
WE MUST CONSIDER	3
WE TEND TO THINK	3
WE WOULD HAVE TO	3
WHAT CAN I DO	3
WILL BE NECESSARY	3
WILL BE USEFUL	3
WOULD BE ABLE TO	3
WOULD LIKE TO THINK	3

CURRICULUM VITAE

PERSONAL INFORMATION

Name : MeltemMuslu
Date/Place of Birth: 24/08/1980, Gaziantep
Email : meltemmuslu@gmail.com

EDUCATION

2009-2014	Çukurova University, Turkey, PhD Institute of Education, Applied Linguistics
2007	Anadolu University, Turkey, MA Institute of Education, Applied Linguistics 'Formative Evaluation of a Process-genre Writing Curriculum at AnadoluUniversity School of Foreign Languages'
2003	Anadolu University, Turkey, BA Education Faculty, Applied Linguistics
1998	James Martin High School, Arlington, TX, USA

WORK EXPERIENCE

2014 Summer	University of Wisconsin, Madison Full-time Lecturer of Turkish Immersion Program
2011 Summer	University of Wisconsin, Madison Full-time Lecturer of Turkish Immersion Program
2010 Summer	University of Wisconsin, Madison Founder of the Beginner Turkish Immersion Program and Full-time Lecturer ,
2008-2009	University of California, Santa Barbara, USA FULBRIGHT Turkish FLTA, Curriculum Programmer

2006-present	University of Gaziantep - School of Foreign Languages, Instructor, Writing Coordinator, Writing Curriculum Programmer, Tester - English Language and Literature Department, Writing Professor
2003-2006	Anadolu University - School of Foreign Languages, Instructor, Writing Coordinator, Writing Curriculum Programmer - Open Education Faculty, Instructor
2003	Ministry of Education, Elementary School-Teacher of English

HONORS and GRANTS

2008-2009	University of California, Santa Barbara. Fulbright Scholar, Turkish FLTA
1997-1998	Exchange Student in Arlington, TX, USA

CONFERENCES, SEMINARS and WORKSHOPS

2013 June	A week-long teacher training course on language teaching College of North West London, London UK
2013 May	25th International Conference on Foreign/Second Language Acquisition, Szczyrk, Poland *'Epistemic Stance Lexical Bundles in L2 Academic Writing: A Contrastive Interlanguage Analysis Study'*
2013 June	Full time teacher training course on language teaching College of North West London, London, UK
2012	Four Skills and Motivating Learners. British Council teacher training seminar
2011	SIELT 2nd Conference, Anadolu University, Eskisehir, Turkey Muslu, M. '*Does Experience Trigger Teacher Burnout?*'
2010	Full-time teacher training course on language teaching

College of North West London

2007 Egypt TESOL, Cairo, Egypt
Muslu, M. *'Feedback: Through the Eyes of the Learners'*

2007 Effective Teaching Seminar
Center of Life-long Learning, Gaziantep University

2006 Symposium of Second Language Writing, Lafayette, IN, USA
Muslu, M. &Saglam, S."*The Role of a Correction Code on Students' Self Correction of Errors in Multiple Drafts Writing in an EFL Context*"

PUBLICATIONS

Bağçeci B. &Muşlu, M (2012). A Study on the Writing Needs of Engineering Students. *Contemporary Online Language Education Journal*: 2(1), 11-21.

Muslu, M., Kutlar, L., Torun, Y., Evisen, N. & Altay, M. (2013) *The Writing Journey.* (ed. BirsenBagceci) EkinYayin, Bursa.