

REFERENCING FOR COHESION IN L2 ACADEMIC
WRITING: A CORPUS ANALYSIS

A MASTER'S THESIS

BY

TIMOTHY P. BENELL

TEACHING ENGLISH AS A FOREIGN LANGUAGE

İHSAN DOĞRAMACI BILKENT UNIVERSITY

ANKARA

MAY 2018

Referencing for Cohesion in L2 Academic Writing

The Graduate School of Education
of
İhsan Doğramacı Bilkent University

by

Timothy P. Benell

In Partial Fulfillment of the Requirements for the Degree of
Master of Arts

in

Teaching English as a Foreign Language

Ankara

May 2018

İHSAN DOĞRAMACI BILKENT UNIVERSITY

GRADUATE SCHOOL OF EDUCATION

Thesis Title: Referencing for Cohesion in L2 Academic Writing

Timothy P. Benell

May 2018

I certify that I have read this thesis and have found that it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Arts in Teaching English as a Foreign Language.

Asst. Prof. Dr. Deniz Ortaçtepe
(Supervisor)

Asst. Prof. Dr. Aysel Sarıcaoğlu,
(TED University - 2nd Supervisor)

I certify that I have read this thesis and have found that it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Arts in Teaching English as a Foreign Language.

Prof. Dr. Julie Aydınlı-Mathews, Ankara Sosyal Bilimler University (Examining
Committee Member)

I certify that I have read this thesis and have found that it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Arts in Teaching English as a Foreign Language.

Asst. Prof. Dr. Zeynep Bilki, TED University (Examining Committee Member)

Approval of the Graduate School of Education

Prof. Dr. Alipaşa Ayas (Director)

ABSTRACT**L2 REFERENCING FOR COHESION IN L2 WRITING:
A CORPUS ANALYSIS**

Timothy P. Benell

M.A. in Teaching English as a Foreign Language

Supervisor: Asst. Prof. Dr. Deniz Ortaçtepe

2nd Supervisor: Asst. Prof. Dr. Aysel Sarıcaoğlu

May 2018

Cohesion in academic and other writing is essential to effective communication. Teachers of English for academic purposes (EAP) place a great deal of emphasis on achieving cohesion, principally through coordinators and subordinating conjunctions. In the classroom and in the academic literature, less attention is paid to the role of referential pronouns to link ideas across clausal boundaries. This corpus-based study compares referencing for cohesion between L1 English writers and L2 English learners. It specifically compares pronominal referencing (*it, she, he, they, them, his, her, hers, its, their, theirs,*) and demonstrative referencing (*this, that, those,*) between two groups in terms of frequency of use, syntactic category, and type of referent, to reveal differences that often undermine the quality of L2 writing.

The L1 English corpus is composed of 383 Economist Leaders articles from 2016 and 2017 (302,618 words) and the L2 English corpus is composed of 371 (388,526 words) essays written by first-year students in an English 101 Composition course in the fall of 2017. Using the corpus analysis software AntConc, concordance searches were produced for all pronominal and demonstrative pronouns and were transferred into Excel sheets for qualitative analysis. Concordance results from either corpus that

included a quotation were excluded from coding since these do not represent original writing. Based on discourse analysis, all pronouns were coded for (a) syntactic function (i.e., demonstrative pronouns, demonstrative adjectives, adjective clauses, noun clauses, and adverbial expressions), (b) part of speech (POS) (i.e., nouns, adjectives), and (c) case (i.e., subjects, objects and complements, idiomatic expressions, and non-referential expressions).

The coded data were analyzed through descriptive and inferential statistics. Raw counts of each pronoun and their percentages were calculated in order to see the overall distribution of occurrences in each corpus. In order to find out whether the differences between the referential occurrences in L1 English corpus and L2 English corpus were statistically significant, means and standard deviations were calculated, and independent samples t-test analyses were run using SPSS.

The findings showed that L2 English writers use referential pronouns differently from L1 English writers. Several statistically significant differences were found between L1 English writers and L2 English learners in referencing including *it* in subject position; *this*, *that*, *these*, and *those* as a demonstrative pronoun, adjective and adverbial expression, *his* and *her* as a possessive adjective as well as *she*, *they*, *them*, and *their*. No statistically significant differences were found between L1 English writers and L2 English learners' use of *theirs*, *hers*, and *him*.

Those major referencing differences observed between L1 English writers and L2 English learners offer some implications for the teaching of academic writing to intermediate students. L1 English writers' use of referential pronouns can serve as a model to academic writing instructors when teaching cohesion in L2 writing. For this, a more detailed qualitative analysis at individual text levels, going beyond analyzing concordance lines as in this study, is required in future research.

Key words: Cohesion, L2 Writing, Corpus analysis, Referencing, Turkish non-native writers

ÖZET

İKİNCİ DİL AKADEMİK YAZMADA BAĞDAŞIKLIK KURMAK İÇİN GÖNDERİM YAPMA: BİR DERLEMBİLİM ANALİZİ

Timothy P. Benell

Yüksek Lisans., Yabancı Dil Olarak İngilizce Öğretimi Programı

Tez Danışmanı: Dr. Öğr. Üyesi Deniz Ortaçtepe

İkinci Tez Danışmanı: Dr. Öğr. Üyesi Aysel Sarıcaoğlu

Mayıs, 2018

Bağdaşıklık akademik ve diğer yazımlarda etkin iletişim için çok önemlidir. İngilizce öğretmenleri yazılı metinlerde bağdaşıklığa oldukça önem vermektedirler, özellikle bağlaçlar yoluyla kurulan bağdaşıklığa. Sınıf içerisinde ve akademik literatürde, bağdaşımsal zamirlerin cümleler arası fikirleri bağlamak için kullanılmasına yeterli önem gösterilmemektedir. Bu derlembilim çalışması ana dili İngilizce olan yazarlar (İngilizce uzman derlemi) ile İngilizceyi yabancı dil olarak öğrenen öğrencilerinin (İngilizce öğrenenler derlemi) bağdaşıklık kurmak için zamir kullanımını karşılaştırmaktadır. Bu iki grup arasındaki zamir kullanımını kullanım sıklığı, sözdizimsel kategorisi, ve gönderim türü gibi genellikle ana dili İngilizce olmayanların yazımlarının niteliğini olumsuz yönde etkileyen farklılıklar şeklinde incelemektedir.

İngilizce uzman derlemi 2016 ve 2017 yılları arasındaki Economist dergisinin Leaders bölümündeki 283 makaleden derlenmiştir ve 302,618 kelimedenden oluşmaktadır. İngilizce öğrenenler derlemi ise 2017 güz döneminde Bilkent Üniversitesi'nde İngilizce 101 Kompozisyon dersini alan birinci sınıf öğrencilerine ait 371 makaleden derlenmiştir ve 388,526 kelimedenden oluşmaktadır. AntConc derlem analiz programını kullanılarak tüm zamirler için dizinler oluşturulmuş, ve bu dizinler

kodlama ve nitel analiz için Excel dosyalarına aktarılmıştır. Hem İngilizce uzman derlemindeki hem de İngilizce öğrenenler derlemindeki direkt alıntı içeren dizinler, yazarın kendi ifadeleri olmadığı için, kodlamadan ve analizden çıkartılmıştır. Tüm zamirler üç kategoride kodlanmıştır: sözdizimsel fonksiyon, sözcük türü, ve cümledeki özne, nesne, tümleç gibi rolleri.

Kodlanan veriler tanımlamalı ve çıkarılmalı istatistik teknikleri kullanılarak nitel olarak analiz edilmiştir. Her bir zamirin iki derlemde de kullanım sayıları ve yüzdeleri hesaplanmıştır. İngilizce uzman derlemi ve İngilizce öğrenenler derlemindeki zamir kullanım farklılıklarının istatistiksel olarak önemli olup olmadığını anlamak için SPSS programı kullanılarak ortalamalar ve standart sapmalar hesaplanmıştır ve bağımsız iki örnek t-testi analizi yapılmıştır.

Sonuçlar, İngilizce öğrenenlerin bağdaşıklık için zamir kullanımının İngilizce uzman derlemindekinden farklı olduğunu göstermiştir. İngilizce öğrenenler ve İngilizce uzmanları arasında zamir kullanımında birçok istatistiksel olarak önemli fark bulunmuştur: *bu, şu, bunlar, şunlar, o, onlar, onların, ve onun. Onunki ve onlarınki* zamirlerinin kullanımında İngilizce uzman derlemi ve İngilizce öğrenenler derlemi arasında istatistiksel olarak anlamlı bir fark bulunmamıştır.

İngilizce öğrenenler ve İngilizce uzmanları arasında bulunan zamir kullanım farkları orta düzey İngilizce öğrencilerine akademik yazmanın öğretilmesi açısından önemli fikirler öne sunmaktadır. İngilizce uzmanlarının zamir kullanımı akademik yazma öğretmenlerine ikinci dil yazmada bağdaşıklık öğretirken model olabilir. Bu modelin geliştirilmesi için ileride yapılacak olan araştırmalar, derlem dizinlerinin analizinin de ötesine giderek, tek tek metin düzeyinde daha detaylı incelemeler ve karşılaştırmalar yapılmalıdır.

Anahtar Kelimeler: Bağdaşıklık, İkinci Dil Yazımı, Derlembilim, Bağdaşıklık, İngilizce Öğrencileri, Writing

ACKNOWLEDGEMENTS

Several individuals deserve recognition for their contribution to making this study possible. First I would like to thank my advisor, Aysel Sarıcaoğlu, who supported me when I was discouraged and opened new paths when I felt my progress was blocked.

I owe a debt of gratitude to my instructors, Necmi Aksit, Julie Aydınli-Mathews and Deniz Ortaçtepe for guiding me on this academic journey.

I would also like to thank the students in my life: first, those in my cohort, especially Aslı, Neşe, Özge, Yaprak, Brent and Yunus Emre whose active participation in class made this academic experience worthwhile; second, the students I have taught over the last 10 years, who have taught me so much about myself and about life.

I am likewise grateful to Tijen Aksit, Director of the Faculty of Academic English (FAE) and Turkum Cankatan, my department head within the FAE, for their special consideration while I pursued my MA TEFL.

Finally, I must thank Özkan Kılıç, who taught me the meaning of “four shoulders”.

TABLE OF CONTENTS

	Page
ABSTRACT	iii
ÖZET	v
ACKNOWLEDGEMENTS	vii
TABLE OF CONTENTS	viii
LIST OF TABLES	xi
LIST OF FIGURES	xiv
CHAPTER 1: INTRODUCTION	1
Introduction	1
Background of the Study	1
Corpus Linguistics and Discourse Analysis	1
Academic Writing: Cohesion	2
Reference	3
Statement of the Problem	3
Research Question	5
Significance of the Study	5
CHAPTER 2: REVIEW OF LITERATURE	7
Introduction	7
State of the Art	7
Discourse Analysis	8
Measuring Writing Quality	9
Cohesion	10

Lexical vs. Grammatical Cohesion.....	10
Cohesion as a Measure of Writing Quality	15
Narrower Studies of L2 Writing and Cohesion.....	15
The Turkish Context.....	16
CHAPTER 3: METHODOLOGY.....	18
Introduction	18
L1 English Corpus.....	18
L2 English Corpus.....	20
Data Analysis	21
Qualitative Analysis: Coding for Syntactic Function, POS, and Case.....	22
Quantitative Analysis	31
CHAPTER 4: DATA ANALYSIS.....	34
Introduction	34
Data Analysis of Corpora Statistics.....	34
It.....	34
Its	39
This	40
That.....	43
These.....	46
Those.....	48
Those as a Demonstrative Adjective	50
Those as a Demonstrative Pronoun	50
They.....	53
Them.....	54
Their.....	55

Theirs	56
Third Person Singular Personal Pronouns	57
CHAPTER 5: CONCLUSIONS	66
Introduction	66
Findings and Discussion.....	66
Use of third person impersonal pronoun <i>it</i>	67
Demonstrative Reference.....	69
Implications for Practice.....	72
Implications for Further Research	76
Limitations of the Study	78
Conclusion.....	79
REFERENCES.....	81

LIST OF TABLES

Table	Page
1 Summary of Corpora Size	21
2 Matrix of Pronouns and Semantic Functions	22
3 Sample Coding of ‘it’ for Cleft or Fronted Structures L1 English Corpus	25
4 Sample Coding of ‘it’ for Idiomatic or Non-referential Structures L1 English Corpus	25
5 Sample Coding of ‘it’ for Unclear References in L2 English Corpus	26
6 Sample Coding of ‘those’ for Type of Reference L1 English Corpus	27
7 Sample Coding of ‘this’ L2 English Corpus	29
8 Sample Coding of ‘that’ in Adverbial Expression L1 English and L2 English	29
9 Sample Coding of ‘z’ in the L2 English Corpus	30
10 Referential Pronouns: Raw Counts and Percentages	31
11 Frequencies of ‘this’, ‘that’, ‘these’, and ‘those’ by Syntactic Function	33
12 Frequencies of ‘it’ by Syntactic Function	34
13 Group Statistics for ‘it’ by Syntactic Function	35
14 T-test Statistics for ‘it’ by Syntactic Function	35
15 L2 Concordance Analysis for ‘it’ in time expressions	37
16 L1 Concordance analysis for ‘it’ in time expressions	37
17 Coding ‘it’ for Unclear References	38
18 Concordance Analysis for ‘its’	39
19 Group statistics for ‘its’	39
20 T-test findings for ‘its’	40
21 Frequencies of ‘this’ by Syntactic Function	40

22	Group Statistics for ‘this’ by Syntactic Function.....	41
23	T-test Findings for ‘this’ by Syntactic Function	41
24	Frequencies of ‘that’ by Syntactic Function	43
25	Group Statistics for ‘that’ by Syntactic Function.....	43
26	T-test findings for ‘that’ by Syntactic Function.....	44
27	Frequencies of ‘these’ by Syntactic Function	46
28	Group Statistics for ‘these’ by Syntactic Function	46
29	T-test findings for ‘these’ by Syntactic Function.....	47
30	Frequencies of ‘those’ by Syntactic Function.....	48
31	Group Statistics for ‘those’ by Syntactic Function	49
32	T-test Findings for ‘those’ by Syntactic Function	49
33	L2 English Writers’ use of ‘those’ as a Non-qualified Nominative Pronoun.....	52
34	L1 English Writers’ Use of ‘those’ as a Qualified Nominative Expression	52
35	Frequencies of ‘they’	54
36	Group Statistics for ‘they’	54
37	T-test Findings for ‘they’	54
38	Frequencies of ‘them’	54
39	Group Statistics for ‘them’	55
40	T-test Findings for ‘them’	55
41	Frequencies of ‘their’	55
42	Group Statistics for ‘their’	56
43	T-test Findings for ‘their’	56
44	Frequencies of ‘theirs’	56
45	Group Statistics for ‘theirs’	57
46	T-test Findings for ‘theirs’	57

47	Frequencies of ‘he’	57
48	Group Statistics for ‘he’	57
49	T-test Findings for ‘he’	58
50	Frequencies of ‘her’ by Syntactic Function	58
51	Group Statistics for ‘her’ by Syntactic Function	59
52	T-test Findings for ‘her’	59
53	Frequencies of ‘him’	60
54	Group Statistics for ‘him’	60
55	T-test Findings for ‘him’	60
56	Frequencies of ‘his’	61
57	Group statistics for ‘his’	61
58	T-test Findings for ‘his’	61
59	Frequencies of ‘hers’	62
60	Group Statistics for ‘hers’	62
61	T-test Findings for ‘hers’	62
62	Frequencies of ‘she’	63
63	Group Statistics for ‘she’	63
64	T-test Findings for ‘she’	63
65	Concordance Analysis for ‘his/her’	64

LIST OF FIGURES

Figure	Page
1 Systems of cohesion in English	13
2 Sample concordance analysis for those in AntConc	22
3 Sample coding part of speech for that in Excel.....	24

CHAPTER 1: INTRODUCTION

Introduction

Corpus Linguistics (CL), briefly defined, is the analysis of a body of authentic language using computer-based querying tools. Discourse Analysis (DA) is often described as the study of connected speech, across the boundaries of clauses and sentences. This study exists at the nexus of CL and DA, examining cohesive patterns in argumentative texts. English writers in academic and persuasive writing use referencing in the form of pronouns to bind a text together and to guide the reader from given to new information. The effective application of this given-to-new paradigm is essential to achieve cohesion across clauses and sentences

Background of the Study

Corpus Linguistics and Discourse Analysis

Over the last 50 years, digital technology has transformed the work of linguistic study. Assembling large amounts of written material into a database for analysis has created a new area of linguistic inquiry called corpus linguistics (CL): the study of language use through automated analysis of collections of transcribed utterances or written texts (McEnery & Hardie, 2015, p. 2). Researchers have exploited CL to analyze discourse to discover patterns in usage among both native speakers (NSs) and non-native speakers (NNSs) (McCarthy & O’Keeffe, 2010). The ability to tag for parts of speech (POS), tense, lemmas and other data within a text has opened up a new area of inquiry which has had an important influence particularly on the study of L2 writing

Academic Writing: Cohesion

Effective academic writing is challenging for any student, but especially for L2 English learners. Since writers often have difficulty in connecting ideas and creating a flow in written discourse, teaching cohesion to learners of English for academic purposes (EAP) requires special attention (Hinkel, 2009). Specifically, the construction of a reader-centric argumentative essay with a unified flow of information is a major challenge for L2 English writers given that argumentation requires the writer to establish clear and unambiguous logical links across clauses, sentences and paragraphs.

Undergraduate EAP programs depend on a fairly consistent formula: the five-paragraph argumentative essay. "The ability to construct supported arguments in English is important for academic success in educational contexts where English is the language of instructions and student assessment is mediated through the academic essay" (Chandrasegaran, 2008). In an effort to acknowledge this significant role of cohesion, researchers have been trying to rationalize the "over 150 classic and recently developed indices related to text cohesion" (Crossley, Kyle, & McNamara, 2015). In order to make sense of the multiple and sometimes ambiguous operational descriptions of cohesion, many researchers have relied on Halliday and Hasan's (1975) description of lexical and grammatical cohesion. Lexical cohesion, which describes the use of vocabulary (synonyms, antonyms, hypernyms, hyponyms, etc.) to connect ideas, is outside the scope of this study. Halliday and Hasan organize grammatical cohesive devices into five categories: reference, substitution, ellipsis, and conjunction.

Reference

Reference ties are divided by Witte and Faigley (1981) into three types: pronominal, demonstrative or definite articles, and comparatives. Biber (1988) indicates that demonstrative pronouns occur frequently in written academic discourse because they build contextual ties between ideas. Among pronominal expressions, personal references may be the least problematic, since they refer to a person, identifiable by sex and number, and likely recently mentioned – or 'presupposed' (Halliday and Hassan, 1975). To achieve cohesion, these references must be endophoric, that is, related to something mentioned in the text. However, managing the scope of *it*, along with demonstrative pronouns *this*, *that*, *these*, and *those*, can be problematic for L2 academic writers (Crosthwaite, 2017). Ineffective deployment of referents may erode the cohesion and coherence of a text if the reader is unable to quickly identify the anaphora.

Statement of the Problem

Although the definition of cohesion is fairly straightforward and concise, the means by which it is achieved are many and varied. The study of cohesion in L2 writing naturally encompasses a vast range variables and measures. Earlier studies sought to measure the relationship between the use of cohesive devices and perception of writing quality (correlation of high-rated/low-rated essays to presence of various cohesive devices). Such analysis has been performed not only on non-native speakers (NNS) (Crossley & McNamara, 2011; Crossley, Kyle & McNamara, 2015) but also on native speakers (NS) (Leńko-Szymańska, 2004; Petch-Tyson, 2009; Witte & Faigley, 1981). Some recent studies have retained this broad approach, looking at reference, substitution, ellipsis and conjunctive/adverbial ties. These studies have sought to understand the frequency and relative frequency of the

range of cohesive devices among distinct English L2 learner populations: Arabic (Aldera, 2016; Hinkel, 2001), Chinese (Crosthwaite, 2017; Yang & Sun, 2012; Liu & Braine, 2005; Hinkel, 2001), Indonesian (Hinkel, 2001), Japanese (Crosthwaite, 2017), Korean (Crosthwaite 2017; Hinkel, 2001), Persian (Zarepour, 2016), Tagalog (Alarcon & Morales, 2011), and Thai (Petchprosert 2013). Some more recent studies have taken a narrower approach, focusing on subordinating and coordinating conjunctions among general ESL populations (Anderson, 2013) and specific populations, including Turkish learners (Yilmaz & Kenan, 2017), Korean learners (Park, 2013) and Chinese learners (Gao, 2016). Only a few studies, however have addressed reference as a cohesive device (Crosthwaite, 2017; Zhang 2015; Naderi, Keong & Latif, 2013; Gray 2010). No studies to date have compared L2 usage with expert usage.

Twenty years ago, Biber, Conrad and Reppen (1998) wrote that "although nearly all discourse studies are based on analysis of actual texts, they are not typically corpus-based investigations: most studies do not use quantitative methods to describe the extent to which different discourse structures are used" (p. 106). Since then, the paucity of such research has been addressed. However, the research that has examined English learners' use of cohesion to maintain texture in writing has been largely restricted to Chinese non-native speakers (Crosthwaite, 2016; Liu & Braine 2005; Ryan, 2015; Zhang, 2014) and European non-native speakers (Leńko-Szymańska, 2004; Petch-Tyson, 2009). Research in Turkish students' use of cohesive ties has been either very broad (Aysu, 2017; Kafes, 2012) or very narrow (Ucar & Yukselir, 2017). Yilmaz and Dikilitas (2017) have explored the conjunctive ties; referential ties remain to be explored. In order to complete the patchwork of

studies on Turkish learners' competence in applying cohesive devices, it is important to address the range of devices in depth.

The syntactic role of the reference in a sentence may also affect the use of the reference such as whether the references are used in subject position or in the object position. However, not many studies have investigated L1 and L2 corpora in the use of referential cohesive devices by employing detailed coding categories of them, considering many possible aspect of their use such as forms and functions. Therefore, there is a great need for such studies to gain thorough insights into interlanguage development of referential cohesion for L2 learners" (Kim, 2012).

Teachers at Bilkent's Faculty of Academic English (FAE) are responsible for improving the writing skills of Bilkent's students. Knowing how our students are using reference may offer insights to our faculty to make these lessons more effective. The writing development of Turkish English for academic purposes (EAP) students may benefit from pedagogical changes informed by the conclusions of the study.

The purpose of this study is to describe how L2 learners use pronominal and demonstrative pronouns for textual cohesion, especially compared to expert users.

Research Question

- How do expert L1 English expert writers and Turkish writers of L2 English differ in their use of pronominal and demonstrative reference?

Significance of the Study

The study is intended to fill a gap in the body of research that examines the role of referential cohesion in L2 writing. Most studies to-date have considered the

use of cohesive devices only within the context of L2 writing. A few studies have compared L2 and L1 performance (Hinkel, 2001; Leńko-Szymańska, 2004; Petch-Tyson, 2009) but none have compared Turkish L1 students' performance with expert usage. The study will also introduce a new metric by comparing how the two groups of writers compare. Knowing to what extent Turkish L1 writers of L2 English use reference ties in the same proportion and in the same way as expert writers can shed light on L1 transfer and improve the quality and effectiveness of writing instruction.

This study will fill a gap in the study of L2 cohesion in English for Academic Purposes by expanding the list of NNS languages to include Turkish. The findings of this study will provide a better understanding of how intermediate level Turkish learners use referencing, which will, as result, offer pedagogical insights unique to Turkish learners.

CHAPTER 2: REVIEW OF LITERATURE

Introduction

Lexicographers and grammarians have relied on corpora, or large collections of authentic language, for hundreds of years to define terms and standardize usage. Over the last 30 years, such collections have become widely available and searchable by means of digital technology, allowing linguists and educators to explore and analyze authentic texts and to unlock their pedagogical potential. During this time, empirical study of lexicography (which includes frequency lists, collocations, functions of words, and meaning of words in context) has dominated the field of corpus research.

The definition of what constitutes a corpus has evolved along with technology. Thanks to the impact of digitization and machine-readable text, the term ‘corpus’ today generally describes not only a collection of authentic language, but one that has been tagged (with metadata) for ease of analysis. According to McEnery, Xiao and Tono (2006) "... there is an increasing consensus that a corpus is a collection of (1) *machine readable* (2) *authentic* texts (including transcripts of spoken data) which is (3) *sampled* to be (4) *representative* of a particular language or language variety” (p. 5).

State of the Art

The value of today’s corpora lies in both their depth and breadth. The largest corpus, the COBUILD corpus published by Collins, contains 4.5 billion words, collected from all manner of print, digital and spoken media (Moon, 2012, p. 197). These huge corpora, such as the Corpus of Contemporary American English,

compiled by Brigham Young University, allow users to restrict queries by genre: newspapers, books, academic journals and even soap opera dialogs. As the genres included expand and the metadata increases, the potential for analysis grows geometrically. Corpora are not limited to the printed word, however. The Michigan Corpus of Spoken English (MICASE) and the British Academic Corpus of Spoken English (BASE) have broadened the scope of linguistic inquiry, allowing researchers to compare a spoken and written English. The wealth of information available about language as a result of digitization has revolutionized the field of linguistics. Although designed as a tool for lexicographers, COBUILD was soon mined for “‘pattern grammar,’ -- explanations of grammatical structures integrated with the specific lexical items most commonly used in in them” (Conrad, 2012, p. 229).

Discourse Analysis

Thornbury (2012, p. 270) calls discourse “both slippery and baggy” meaning that “it eludes definition” and “embraces a wide range of linguistic and social phenomena”. Its synonym might be communication, whether spoken or written. He states that discourse can vary depending on context or “the describable internal relationships” (p. 270). Nunan (1993) describes discourse simply as a “communicative event” and discourse analysis (DA) as “the interpretation of the communicative event in context” (p. 6-7). Such analysis is most often and most easily applied to texts, which may comprise significantly smaller collections of thousands of words, rather than hundreds of thousands or millions of words. Corpus Analysis (CA) has facilitated the linguistic analysis of not only lexical items, but also of grammar, syntax and of discourse.

The literature describes a ‘cultural divide’ between corpus linguistics and discourse analysis, which, at a basic level, is the difference between quantitative and

qualitative analysis (Flowerdew, 2012, p. 84). However, Biber, Conrad and Reppen (1998) urge that “corpus-based analyses must go beyond simple counts of linguistic features” and that they ought to include “qualitative, functional interpretations of quantitative patterns” (p. 5). In fact, corpus linguistics has been applied to facilitate discourse analysis, which examines patterns of usage (McCarthy & O’Keeffe, 2010). Thornbury (2010) concludes that “corpus discourse analysis must then, by definition, avail itself of quantitative methods with the aim of producing findings that are both descriptive and explanatory” (p. 270).

McEnery et al. (as cited in Flowerdew, 2012) have suggested that the cultural divide between CL and DA may be narrowing and that the two approaches may in fact be complementary. Flowerdew (2012) asserts that from the perspective of written corpora, the aims of CL and DA may overlap: “This complex synergy of fields, with corpus linguistics no longer hovering on the periphery of discourse analysis but now assuming a central role” (p. 110). This synergy has created new challenges “for both software developers and corpus analysts” (p. 110). This study may straddle these classifications.

Measuring Writing Quality

The use of text and CA to discover patterns and correlations in L2 writing has focused on argumentative writing in an EAP environment (Chandrasegaran 2008; Leńko-Szymańska, 2004; Petch-Tyson, 2009; Zhang 2015). Within this context, researchers have looked at a number features that determine writing quality including clause types (Becker et al., 2016) discourse structures (claim and support) (Stab & Gurevych, 2014), qualification, and certainty (Hyland & Milton, 1997)

Cohesion

Hinkel (2009, p. 279) defines cohesion as “the connectivity of ideas in discourse and sentences to one another in a text, thus creating the flow of information in a unified way”. Halliday and Hasan (1976) define cohesion as the “the relations between two or more elements in a text that are independent of the structure (p. vii). To achieve cohesion in written discourse, references must be endophoric, that is, related to something mentioned in the text. Exophoric references, often used in speech where there is a shared visual or other sensory experience, are not relevant in written discourse. A cohesive text is one that effectively bridges the 'given' and the 'new' information across clauses and sentences, allowing the reader to follow the writer, making clear the connections between persons, objects or concepts. In a cohesive text, the references among elements are described as “recoverable”; that is, their connections are readily apparent. To study cohesion, then, is to identify what distinguishes a text from a disconnected sequence of sentences.

Lexical vs. Grammatical Cohesion

Halliday and Hasan (1976) categorized cohesion into two types: lexical and grammatical. Lexical cohesion describes the use of repetition, synonyms, hyponyms, hypernyms and other related words to represent a previously stated word or idea. DA focusing on lexical cohesion demands painstaking manual examination of a text and does not lend itself well to CA. Studies of lexical cohesion of academic texts have typically relied on sample sizes of 20-100 essays (Güngör & Uysal, 2016; Kafes, 2012; Liu & Braine, 2015; Park, 2013; Zarepour, 2016).

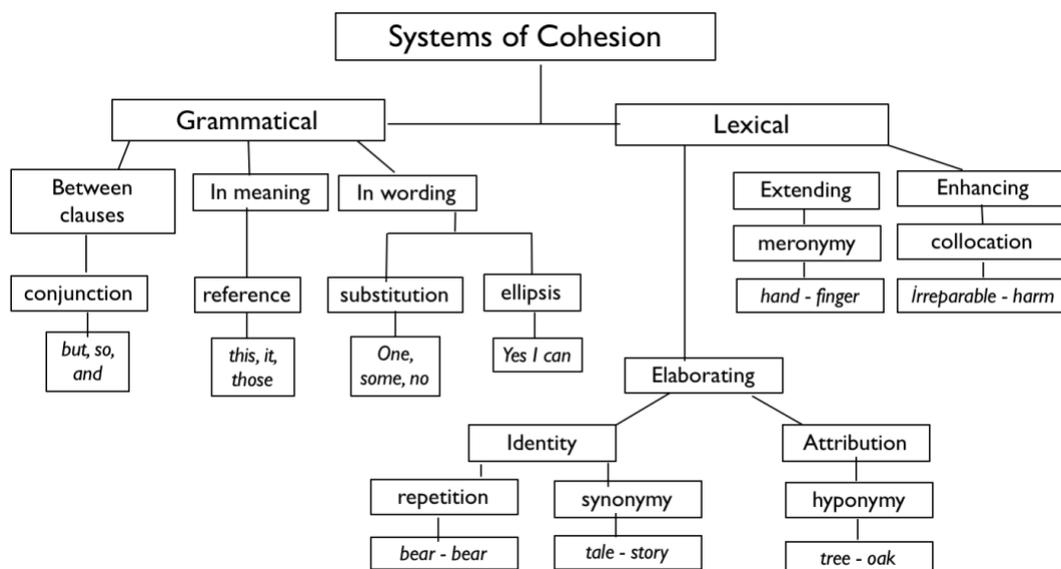
Grammatical cohesion connects ideas in a text by means of reference, substitution, ellipsis and conjunction (Halliday & Hasan, 1976). In a text, the terms 'reference' and 'refer' describe the function of words like pronouns, determiners, and

demonstratives to designate a noun phrase, an argument, process or an event that they identify within the immediate co-text. *Anaphoric reference* describes backwards referencing (John is a student, but he is also an athlete); *cataphoric reference* describes forward referencing (This may shock you, but Mrs. Clinton won the popular vote). When a referent is too far away from the antecedent, or unsuccessfully represents the lexical item that defines the reference, then cohesion is said to be broken, and the writing loses its 'texture' (Halliday & Hasan, 1976).

Personal references include personal pronouns, possessive determiners (also called possessive adjectives) and possessive pronouns. With the important exception of *its*, these may be the least problematic for L2 learners, since they refer to a person (or non-humans), identifiable by sex and number and likely recently mentioned or “presupposed” in the text. Non-cohesion occurs when the reference is ambiguous. “Mary had a cat and a dog but it died”. Here, the reference to anaphoric reference is unclear: it could be either the cat or the dog. In the expression “Mary and her mother had a cat and a dog, but they died”, the reference is also ambiguous, although the reader is likely to connect the personal pronoun to the pets, since owners typically outlive their pets, and because the antecedent are physically closer to the pronoun. This is to say that reader brings some expectations in order to decode a text. These expectations help the reader resolve referents in expressions such as “Mary adopted a new cat. Her mother must be happy about it”. In this case the possessive pronoun *her* clearly refers to Mary. The object pronoun *it* is more difficult to resolve: it could refer either to the pet itself, or to the action (Mary’s having adopted the cat). In either case, the meaning is clear enough and cohesion is achieved. However, when reader expectations are insufficient to decode a text, it results in a lack of cohesion.

A second category of reference is demonstrative reference. Halliday and Hasan (1976) describe demonstrative reference as a kind of “verbal pointing” on a scale of proximity. These pronouns include *this*, *that*, *these*, and *those*. Brown and Yule (1996) simplify the discussion of grammatical cohesion through reference in text and discourse with as follows:

- a. Repeated form: The Prime Minister recorded her thanks to the Foreign Secretary. The Prime Minister was most eloquent.
 - b. Partially repeated form: Dr E. C. R. Reeve chaired the meeting. Dr Reeve invited Mr Phillips to report on the state of the gardens.
 - c. Lexical replacement: Ro's daughter is ill again. The child is hardly ever well.
 - d. Pronominal form: Ro said she would have to take Sophie to the doctor.
 - e. Substituted form: Jules has a birthday next month. Elspeth has one too.
 - f. Elided form: Jules has a birthday next month. Elspeth has too.
- (p. 193)



Adapted from Halliday and Hassan, 1976

Figure 1. Systems of cohesion in English

In the 40 years since it was first published, Halliday and Hassan's taxonomy of cohesion has served as the basis for analysis of L2 writing in English. The authors describe two types of cohesion: grammatical and lexical. Simply stated, lexical cohesion depends on a level of specific vocabulary (repetition, synonymy, hyponymy, meronymy and collocation) to carry meaning across a text, while grammatical cohesion depends on a range of classes of functional terms (linkers and conjunctions, pronouns, substitution words and ellipsis). (see figure 1) This framework forms the basis of a number of L2 studies, in both spoken and written language.

The studies of cohesion in ELL writing have been both broad and narrow in scope. The broader studies have chosen any number of the elements from Halliday and Hassan's (1976) framework to describe the use of cohesive differences within one L2 English writing population (Alarcon & Morales, 2011; Aysu, 2017; Zarepour, 2016; Zhang 2015), among diverse groups of L2 English writers (Hinkel 2001; Crosthwaite, 2017), between L1 and L2 English writer populations (Crossley

& McNamara, 2009; Liu & Braine 2004; Petchprasert, 2013), between non-native writers of different levels (e.g., Leńko-Szymańska 2009; Yang & Sun, 2011) or looked for correlation between writing scores and the occurrence of cohesive elements (e.g., Witte & Faigley, 2017). On the whole, these studies suggest that cohesion is a key element of argumentative writing, and that L2 learners face challenges in effectively applying cohesive devices.

Broader Studies of L2 Writing and Cohesion

Despite the nearly universal application of Halliday and Hasan's (1975) taxonomy, comparisons are not always easy. The variety of devices has allowed researchers to cherry-pick elements, making comparisons between studies difficult. Still, there are still some bases for comparison. Liu and Braine (2005), in a text analysis of 50 student essays, discovered that Chinese undergraduate writers depend most on lexical devices, referencing, and linkers, in that order. The study discussed the problems with lexical cohesion, including an overdependence on repetition. A review of the research literature from the 1990s to the early 2000s seems to confirm that L2 English writers overused repetition of the same lexical item to achieve coherence across a text (Liu & Braine, 2005).

Liu and Braine (2005) also reported inconsistent use of pronouns, such as shifts from plural to singular, and from second to third person. In this study, "the quality of writing was also revealed to significantly co-vary with the number of lexical devices and the total number of cohesive devices used" (p.623). In a study of L2 academic writer using similar methodology, Alarcon and Morales (2001) reported results that differed from Liu and Braines: among Tagalog speakers, reference accounted for 90% of total cohesive devices in their study.

Zarepour (2016) analyzed the writing of Iranian EFL learners to discover referencing was the most frequently used form of cohesion (43% of all occurrences) and similarly, represented 43% of all cohesive errors. The study also showed that lexical cohesion was the second most commonly deployed tool, most of which was repetition. Zarepour also calculated errors, the most frequent of which were related to reference, followed by conjunction, lexical cohesion, ellipsis, and substitution. In the case of reference cohesion, major portions of errors were related to personal pronoun and demonstrative pronoun.

Cohesion as a Measure of Writing Quality

Using corpus analysis to study cohesion presents a set of problems. Query tools cannot easily identify cohesive elements in a text such as, pronoun referents, substitutions, and ellipsis. To identify what makes individual text cohesive requires manual text analysis. Given the multiple functions of the referent pronouns in English, corpus analysis may involve manual disambiguation of multi-use terms such as *this, that, these, those, her, and his*.

From the 1990s to the early 2000s, a large number of studies were published which sought to find a correlation between the use of cohesive devices and writing quality (based on scores). Most studies, including from Alarcon and Morales (2011), found no correlation between frequency of cohesive devices and writing quality, possibly because the mere presence of a cohesive device did not mean that it was effectively executed.

Narrower Studies of L2 Writing and Cohesion

The narrower studies, those which focused on a single element of cohesion, have for the most part considered elements of grammatical cohesion, whether linking (adverbials and coordinators), or pronominal.

Managing the scope of *it* and *this* is particularly problematic in L2 academic writing (Hinkel, 2001; Kim, 2012; Swierzbis, 2010), leading to a serious loss of coherence for a text if not appropriately managed. This loss of coherence is the result of the reader's inability to retrieve previously-mentioned information with the level of accessibility encoded by the referring expression (e.g. high accessibility for *it*, but mid-accessibility for *that*). Kim (2012) found that Korean EFL writers overused *it* when referring to long sequences of text where the demonstrative pronoun would be appropriate in the L2 target, while Hinkel (2001) found that Korean EFL writers frequently produced demonstrative pronouns that did not clearly relate to a given referent in text. Likewise, a study comparing American L1 English writers to L2 English writers of Dutch, French, and Finnish backgrounds found "the intended referent is obscured by the choice of a referring expression which is either insufficiently specified...the problems were mostly related to sloppy use of *this*" (Petch-Tyson, 2009). Likewise, Leńko-Szymańska (2004) reported that Polish learners of varying levels use *this*, *that*, *these* and *those* at statistically significantly higher rates than L1 English writers.

Whether these L2 patterns exist in Turkish writing has been heretofore unknown.

The Turkish Context

Of the few studies that examined cohesion in the writing of Turkish learners of Academic English, Aysu (2017) analyzed the use of discourse markers among elementary-level prep students. The study showed among the 180 discourse markers used, more than 50% were 'and', and more than 25% were 'but', indicating a very limited range. Öztürk & Köse (2016) compared the frequency of lexical bundles as linkers in PhD level writing among Turkish and native English speakers. The

research indicated that use such bundles much more frequently but use a much smaller range. Kafes (2012) studied Turkish EFL learners' ability to compose cohesive texts in their first language and in English, to learn whether similarities existed between lexical cohesive ties. The author found that repetition accounted for more than 70% of the lexical cohesive ties in English L2 writing of Turkish students (and 55% of lexical cohesion in their L1). Still, there has been no study to-date that explicitly examines any of the range of grammatical cohesive devices used by Turkish university-level writers.

This study will contribute to the literature by providing a detailed analysis of the use of pronominal referencing, the most frequent type of referencing in L2 writing (Hinkel, 2001). The study will also represent the first large-scale ($n > 30$) study of pronominal referencing by Turkish students of L2 English using corpus analysis techniques combined with qualitative discourse analysis.

CHAPTER 3: METHODOLOGY

Introduction

Given that the literature explains the challenges non-native speakers (NNS) face in achieving cohesion in their texts, and the high frequency of referencing to achieve it, this study will compare the use of native speaker (NS) and (NNS) use of pronominal and demonstrative referencing for cohesion in the genre of persuasive or argumentative writing. To this end, two corpora were created. The first corpus represents NS writing and is hereafter referred to as the L1 English corpus, (i.e., English as a first language corpus or expert English corpus). The second corpus contains NNS writing of L1 Turkish students in the persuasive or argumentative style and is hereafter referred to as the L2 English corpus, (i.e., English as a foreign language corpus or student writing corpus).

L1 English Corpus

The L1 English texts which compose the corpus represent two years' worth of Leaders articles (2016-2017) from the Economist, a total of 384 articles and 302,618 words. These pieces of writing were chosen as the model L1 English corpus for several reasons. First, the Economist newspaper's commitment to a clear, uncomplicated and direct writing style is sound model for academic writers. "The first requirement of The Economist is that it should be readily understandable. Clear writing is the key to clear thinking. So think what you want to say, then say it as simply as possible" (The Economist Style Guide, p.1). The Style Guide, quoting from *Fowler's Modern English Usage*, further emphasizes the importance the paragraphs of "a unit of thought, not of length" (p. 3). As such, the goals first-year

academic writers are reflected in the style guide of the Economist. Beyond style, the rhetorical purpose of the writing in the two corpora is the same: to persuade. The English corpus is composed of articles only from the “Leaders” section, the publication’s opinion and editorial section. The decision to build the L1 English corpus from opinion pieces is intended to maintain an “apples-to-apples” comparison of argumentative or persuasive writing, wherein writers introduce a thesis or recommendation supported by evidence.

The essays of first year students and opinion/editorial writing both engage the reader with controversial topics, presenting evidence in an attempt to convince the reader of the merits of a particular point of view. While the styles of the two corpora are not identical – The Economist is often written in a cheeky style and includes fanciful vocabulary such as “hotch-potch” and “shindig” – any effect of these differences will not be relevant to the analysis that is focused on cohesion and referencing. In this way there is a clear logical correspondence between the genres of writing in the two corpora.

The L1 English corpus of Leaders articles, the complete set of opinion writing from 2016 and 2017, were downloaded in PDF format from the website of The Economist and converted to .txt format. Titles, subtitles, promotional copy and other extraneous text were removed or otherwise excluded from analysis. One limit of the correspondence between the two corpora is the Leaders section’s focus on news events. This focus has the effect of including more time references (this week, this month) which do not occur in L2 academic prose. For this reason, such time references are excluded from the analysis.

L2 English Corpus

The L2 English corpus is composed of 371 essays and 388,526 words written by Bilkent University students for their English 101 writing course in the fall of 2017, and submitted electronically, usually in MS Word format. The essays represent the work of roughly 150 unique students (or 8% of the total 1,860 students enrolled in English101) from nine sections (from a total of 93 sections) taught by nine instructors (two male and seven female), including both native Turkish and native English speakers.

For this research project, which received approval from the Ethics Committee of Bilkent University, texts were collected from English 101 instructors who, after being informed of the nature of this study, agreed to share their students' work. Students whose work is included in this study gave their consent to anonymously participate in this study for research purposes. For the most part, the students in English 101 are first-year students who have recently passed the exit examination of the Academic English Preparatory School of Bilkent University's School of English Language (BUSEL). It is possible that the sample contained repeat students, although such students usually take the spring semester course.

The aim of the freshman English course, English 101, is to “[introduce] students to an academic approach to thinking, reading, speaking, writing and language use; skills they will need in their departmental studies. The course also aims to develop students' linguistic accuracy and range in English” (Bilkent University Faculty of Academic English). The Bilkent English101 course is a 14-week content-based English freshman-level composition course in which students develop their critical thinking and academic writing skills. It is the first credit-bearing course in English that the students must take. During the 14-week semester,

English101 students must write three argumentative-style essays. These essays must refer to a reading list selected by the instructor. Although course content and readings for the nine sections in this sample are unique, the grading criteria for these essays are standardized.

Among the general themes in the fall 2018 semester are racism, human intelligence, psychopathy, and religious freedom. The essays follow the standard five-paragraph model with a thesis and topic sentences supported with information from assigned class readings. The prompts included the following questions: Is Artificial Intelligence a Gift or a Threat? Should Autism be Cured? Is Psychopathy a Matter of Nature or Nurture? Should Genetic Engineering of Human Intelligence be Permitted? L2 English essays were likewise converted either from .pdf or MS Word format to .txt format, anonymized and stripped of bibliographic information.

Table 1

Summary of Corpora Size

	L1 English Corpus	L2 English Corpus
No. of Files	383	371
No. of Words	302,613	388,526

Data Analysis

Any identifying information or other text unrelated to discourse was manually removed from the raw text files. In the case of student essays, this meant works cited information and student name, section number and date of submission. For The Economist articles, this meant interstitial promotion, subheadings, and web links to other articles. The cleaned writing samples were saved in individual files and copied into separate directories to create two corpora.

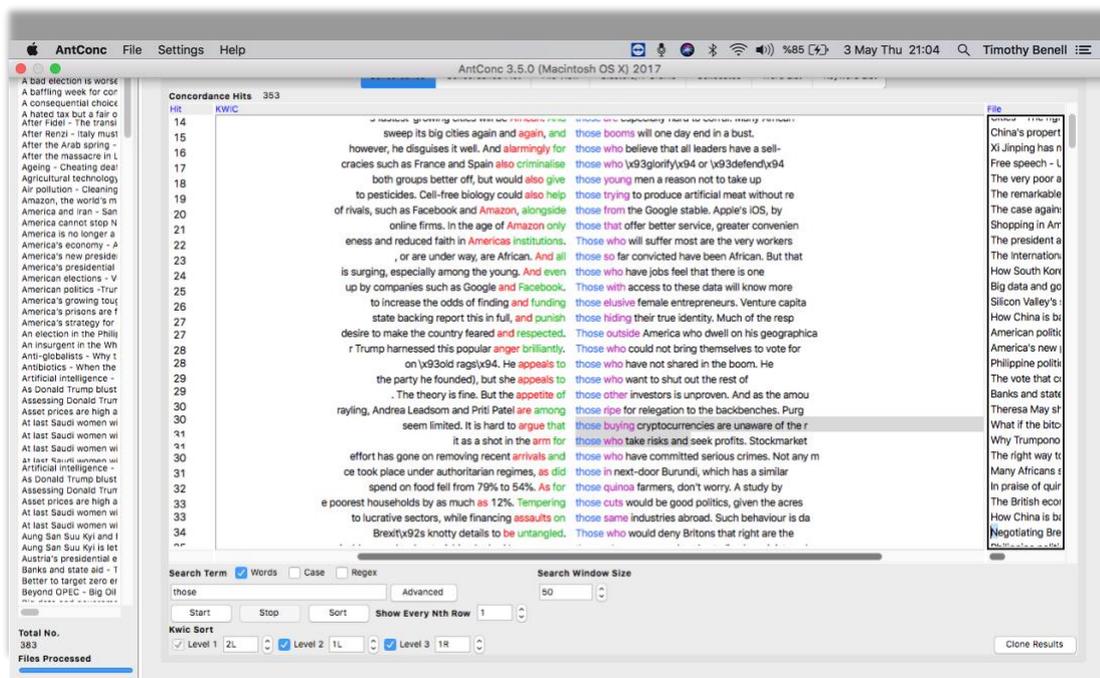


Figure 2. Sample concordance analysis for *those* in AntConc

The next step was to load the corpora into AntConc, a freely-downloadable corpus analysis software¹. Concordance searches were produced for relative pronouns *it*, *its*, *they*, *their*, *theirs*, *this*, *these*, *that*, *those*, *he*, *his*, *him*, *she*, and *her*. Results of each query were exported into text files using AntConc's export function. These text files were then imported into an Excel file, one for the L1 English results and one for the L2 English results. In each file, tabs were created to hold the results of each of the concordance results by pronouns listed above.

Qualitative Analysis: Coding for Syntactic Function, POS, and Case

Table 2

¹ After loading files into AntConc, some extraneous text was sometimes found in the text files. This text was removed from the files but did not affect analysis.

Matrix of Pronouns and Semantic Functions

		Pronouns															
		it	its	this	that	these	those	they	them	theirs	he	her	him	his	hers	she	his/her
Semantic Function	Nominative Expressions																
	Nominative Pronoun			X	X	X	X	X	X			X		X			
	Argument																
	subject	X		X	X	X	X	X		X	X			X	X	X	
	object	X		X	X	X	X		X	X		X	X	X	X		
	complement	X		X	X	X	X	X		X	X			X	X	X	
	unclear reference	X															
	Cleft or Fronted Structure																
	cleft subject	X															
	cleft object	X															
	Non-referential/Idiomatic	X															
	Time Expressions	X		X	X	X											
	Noun clause				X												
	Generic Reference																
	qualified (non personal)							X									
	qualified reference to people							X									
	Adjective Expressions																
	pos/demo adjective					X	X					X		X			X
	adjective clause				X												
	relative clause (non-time)				X												
	Adverbial Expressions			X	X												

Pronouns were first coded for part of speech. In the case of *he, she, it, they, them, him, hers* and *theirs*, this coding was not necessary since these pronouns have fixed and unambiguous syntactic roles as nominative pronouns. The same was true for *his, its* and *their* which function only as possessive adjectives. Multi-function pronouns *that, these, this, and those* (POS) were coded as either demonstrative pronouns (dp) or demonstrative adjectives (da). Further coding for *that* was necessary to separate adjective clauses (ac) as well as noun clauses (nc). Adverbial expressions were also coded separately. Although this study is mainly concerned with referencing, the resulting data on the proportional deployment of these terms in their multiple functions is also of interest.

	B	C	D	H
	text preceding anaphora	POS	Code	end text
1304	. Does black people tend to commit crime? Or is it just a negative discrimination, a bias	ac		that we learned from various sources? Egalitarian aspect can lead us to right answer. They ar
1305	ending to solve their problems without government involvement but as if some facts from the world	ac		that we live considered, it seems to be an utopic idea. It may sound liberal but in
1306	the social construction it causes. Ever since we're born, we engage with the culture	ac		that we live in. Culture determines what we do, how we should behave, how we see the
1307	anti hero this help audience to sympathize with anti heroes. Another idea to support preceding one	ac		that we may find satisfaction with anti heroes' spiteful decisions because our behaviour is det
1308	ware of how dangerous the social media is. Pedophilia and perverts are the greatest enemies	ac		that we need to protect ourselves and our children in our daily life. But today they are
1309	religion can make with it. If we all respect each other and our beliefs, the society	ac		that we part of it, become habitable. Media has a huge and inevitable impact on people abo
1310	n the industry are right about being anxious. Algorithms are complicated. They are so complicated	adv		that engineers might not know what they will turn out to be until they activate them. That
1311	rt and motivation and engagement with education. (Collin, 2011). There are so many online courses	adv		that a high school student can take to improve his or her ability to get into a
1312	images brain and causes antisocial behaviors. When children grow up in good environment the gene	adv		that active does not affect however when child is abused this gene is activated and cause som
1672	s has caused heated controversy, especially in Europe. (Langford, Full face veil) We can encounter	da		that unfortunate occurrences, particularly in Western Europe countries. The reason for that is
1673	ical neglect" in the United States. (Swan, 2006) According to these rates, many children died from	da		that unreasonable reason and it is likely that if the government will not intervene that issue sc
1674	than 41 million. In World War 2 atom bomb was used and over 60 million people were killed in	da		that war. With the developed technology the casualties in wars increased over years. Scientist
1675	exchanging ideas and coming up with plans in terms of helping people. People will benefit more	da		that way and the outcome will be more pleasant for the both sides who participates in the
1676	in commercials, you will start to think that all people's life has to be	da		that way. For instance, in food and drink commercials, women always bring the food or the dr
2073	th an evil character or some character with less morality. Wilterdink (2015) gives an example about	dp	obj	that with Game of Thrones' characters which is there are some evil characters seem unlikable i
2074	more sensitive about the intake of the type of food that enter their body. Added to	dp	obj	that, with taking advantage of the developments happening in the world, it is now an option t
2075	world is making people sick and it seems that this situation will continue to be like	dp	comp	that with the presence of major capitalist forces playing with the health and lives of people.
2076	services for some people whose religions which is limited. Bouchard(2017) has a good example about	dp	obj	that. He suggests: "In winter, if a woman is wearing a full veil with her two
2077	ants meaningful roles called "teacher" and "learner". (Millgram Experiment (video) 2012)	dp	subj	That helps people to fully get in that role because a teacher is meant to teach not
2078	"charisma is a key value for many anti heroes" according to Jason Mittell (Wilterdink, 2015).	dp	subj	That helps us to see anti heroes' heroic ways. Audience may no more think about anti
2092	a black slave and this shouldn't be tolerated. Can you imagine the psychological trauma	nc		that a African American person living at that time? They must have felt really humiliated. Euroj
2093	are powerful enough to make people end their own lives. The second issue to consider is	nc		that a big lie can form "tunnel vision" in humans. That means when a big
2094	over Black people are reflected on each part of cultural dimensions which society owns. The fact	nc		that a black mother was painted over white and her children were called "nigger" by the white
2095	they immediately popped up with the idea that the person was black as they could predict	nc		that a black person could do what has been mentioned in the experiment, that shows that the
2096	among women. People are able to win this racist approach easily because the advertisement show	nc		that a black woman take of her t-shirt after using Dove product then while taking of
2097	nor of Ned Stark", he ordered 8th favorite character in polls. (2013) This example clearly shows	nc		that a character's charm and charisma can make him favorable and popular among the audien

Figure 3. Sample coding part of speech for *that* in Excel

The POS coding task for the demonstratives represented 15,552 judgments. Once the nominative functions were identified, the next task was to determine whether these references functioned as subjects, objects or complements. Given the unambiguous role of 'him' and *them* as object pronouns, this task required 20,075 judgments. Finally, the nominative pronouns were judged to be anaphoric or anaphoric references, for another 20,446 judgments for all nominative pronouns, personal and demonstrative. In the end, less than 1% of references were judged to be cataphoric. These results were judged to be not meaningful and were excluded from the reported results.

The last phase of coding was concerned with sorting the idiomatic uses of two pronouns: *it* and *those*. In addition to representing a singular non-personal referent, the pronoun *it* can be used in a cleft or fronted structure:

Table 3

Sample Coding of 'it' for Cleft or Fronted Structures L1 English Corpus

Preceding text	POS	Arg	Type	end text with reference	Source file
The country has a strong and longstanding commitment to democracy, and Colombian voters have shown no liking for Marxists.	p	s	<u>cleft</u>	<u>It</u> will take a generation, genuine contrition and an ideological conversion for the FARC t	Ending Latin America's oldest war - A messy but necessary peace.txt
South Korea and Taiwan have enjoyed strong manufacturing output and exports on the back of a reviving world economy. But	p	s	<u>cleft</u>	<u>it</u> is hard to feel upbeat about the prospects of such export-leaning economies if	Investing in emerging markets - Turkeys and blockbusters.txt
Industrial plants are shutting down. Unemployment is high. In such poverty traps	p	s	<u>cleft</u>	<u>it</u> is easy to misconstrue free-trade deals as giving supranational capital the right to	Trade agreements - Asterix in Belgium.txt

It is also possible for *it* to behave in an idiomatic or non-referential way, or as an expression of time as in the following examples:

Table 4

Sample Coding of 'it' for Idiomatic or Non-referential Structures L1 English Corpus

Preceding text	POS	Arg	Type	Ending text with reference
served notice that it wanted to leave the overbearing, unrepresentative union to which it had long been shackled. And so	p	s	<u>idiomatic</u>	it was, -- but not in quite the way that Theresa May had imagined. Britain's
The world must do what it can to thwart such plots, though some will doubtless succeed.	p	s	<u>idiomatic</u>	It is worth recalling that America has been here before.
(who even won a Supreme Court case forbidding the government from triggering Brexit without Parliament's permission),	p	s	<u>non-referential</u>	it at last looks as if independence beckons. This week
Does anyone seriously imagine that this power would not be abused?	p	s	<u>non-referential</u>	It is as if Mr Sarkozy wants to turn a drunken rugby chant

Table 4 (cont'd)

The less he can impose his version of xenophobia and Euroscepticism on the Netherlands the better. Unfortunately, however,	p	s	<u>time</u>	it is too soon to celebrate the rollback of populism. The very idea of a
The bull market in everything asset prices are high across the board. Is	p	s	<u>time</u>	it time to worry? With ultra-loose monetary policy coming to an end,
After years of falling prices and fitful growth, Japan's nominal GDP was roughly the same in 2015 as	p	s	<u>time</u>	it was 20 years earlier. America 's grew by 134% in the same time period;

Occurrences of *it* in the L2 English corpus were also coded for unclear references. Although the reference intended by the writer became clearer after several checks, there remained some persistent questions regarding the referent. It is possible that even closer reading may reveal the intended referent, or that these may represent failed idiomatic or non-referential expressions. Some examples follow:

Table 5

Sample Coding of 'it' for Unclear References in L2 English Corpus

Preceding text	POS	Arg.	Type	Ending text with reference	Source file
preferred women who would laugh at their jokes to those who made jokes. Women, however, preferred partners who were funny...."(2). In other words, as society thinks women likes men who are funny,	p	s	<u>unclear reference</u>	it is above the rumour and that is more deeper. However, the person who makes the jokes must be men. In addition, "	T7E3S (1).txt
. Another example, children can be jealous and they can harm their friends toys or bodies, but, it does not make them evil because they are not able to think what will be happened.	p	s	<u>unclear reference</u>	It is another situation to act like a demon because it is not possible to predict what is bad or good and it can creates bad situations, so these actions can be bad. Conditions	T5E3S (26).txt
he society and people or cause people to lose their rights ,government should intervene in these religious differences to some extent. This intervention shouldn't be more than necessity because if	p	s	<u>unclear reference</u>	it is more this cause people to lose their religious rights and freedoms and exclusion of people who have different religious beliefs in society. Everyone has freedom of religion and belief but i	T6E3S (7).txt

Finally, *those* can function as a demonstrative adjective (da) or as a demonstrative pronoun (dp). As a dp, *those* often functions as a specialized reference to third persons or objects (plural), often with a relative clause or reduced relative clause attached. This use of *those* + qualifying phrase was considered worthy of special coding as a dp with a qualifying phrase attached, and was distinguished from general anaphoric references.

Table 6

Sample Coding of 'those' for Type of Reference L1 English Corpus

Preceding text	POS	Arg	Type	Ending text with reference	Source file
that what matters most is what happens in the classroom. The successful children are	dp	c	qualified ref persons	<u>those</u> who are exposed to good teaching more often. Having pupils turn up is a	Homework for all - What countries can learn from PISA tests.txt
Democratic Party. America is not alone. Across Europe, the politicians with momentum are	dp	c	qualified ref persons	<u>those</u> who argue that the world is a nasty, threatening place, and that wise nations	Globalisation and politics - The new political divide.txt
compared with what Mr Trump proposes. On plenty of other questions her policies are	dp	c	qualified reference non-persons	<u>those</u> of the pragmatic centre of the Democratic Party. She wants to lock up fewer	The presidential election – America's best hope.txt
the time, plebiscites lead to bad politics and bad policy. The most problematic are	dp	c	qualified reference non-persons	<u>those</u> on propositions that voters do not understand or subjects	The referendum craze - Let the people fail to decide.txt
Modern food also involves more nutrients and vitamins than	dp	o	qualified reference non-persons	those found in traditionally grown food. Being able to gain adequate nutrients	T4E3S (29).txt
some similarities between psychopaths behaviours and	dp	o	qualified reference non-persons	those of children. This theory can prove that genetic roots may play role in the existence of the psychopathy	T9E1S (11).txt

Table 6 (cont'd)

Although many options to provide healthier options of food are being developed,	dp	s	Non-qualified reference	those are not available to a large share of the world population yet.	T4E3S (20).txt
The disadvantages should not lead people to stop using it. Even though	dp	s	Non-qualified reference	those are real facts, the advantages on the other hand makes social media	T3E3S (30).txt
marking him as a "abnormal" person which called "others", makes character at least exceptable.	dp	s	Non-qualified reference	Those are the reasons of decent rise of anti-heros and moralities part of effect audience is hugh.	T8E3S (7).txt
nobody should stop using social media, because while it may have risks for adolescents,	dp	s	Non-qualified reference	those could be solved with proper education to both parents and adolescents,	T3E3S (14).txt

Excluded from these concordance results were any occurrences which figured in a quotation from either corpus since these do not represent original writing.

Judgments regarding usage overlooked surface level grammatical errors. For example, *it's* used as a possessive adjective was coded as a possessive adjective, despite the spelling. Ungrammatical referencing was also ignored for the sake of this analysis. For example, if *it* referred to a plural head noun, it was still included in the analysis. Instances of *that* in poorly constructed noun clauses were still coded as noun clauses; for example, "Today, it is not clear that who the strangers are...". Lines of corpus software generated instances were reviewed multiple times. In each round of coding, the clearest cases were coded and the less clear cases were isolated for later review.

Below is a sample from the L2 concordance for the word *this*. Columns were added to code for part of speech and argument. The original intention was to identify the type of referent, whether a simple reference (nouns, noun phrases) or an extended reference (clauses, sentences, or larger concepts). However, this analysis was beyond the scope of this study and is recommended for further inquiry. The last column

shows the source file for the concordance. In the first case, it is source file is the third essay (E3) from the eighth instructor (T8) from student 24.

Table 7

Sample Coding of 'this' L2 English Corpus

Preceding text	POS	Arg.	Ending text with reference	Source file
and some have argued that they may include political events such as 9/11, which gave people the desire to search for vigilante figures, even though	dp	s	this appears inconvincing. Other arguments include the effect of charisma, moral alignment and to suggest that morality is a factor th	T8E3S (24).txt
this double-edged sword is sharper appears to be a simple task. The battle in todays world is always between science and ignorance, as we have seen	dp	s	this applies to our modern nutritional trends as well. Some public figures do their best to help us discover what is objectively good for us,	T4E3S (11).txt
these actions are ignored and the audience will have a connection with them. Some of the reasons for	dp	o	this are charisma and charm, fascination, motivation and relative morality of the anti-hero and his/her actions (T8E3S (4).txt
is involved with business we automatically think about a man rather than a woman. Why? Because social engineers lead us to have	dp	o	this as a prejudice. The point of this essay is media's role on creating	T2E3S (7).txt
none of them contributing to military services is unacceptable since other citizens see	dp	o	this as a social inequality. The contribution to the workforce of the Haredim is nowhere near the ordinary citizen	T6E3S (2).txt

Uses of *this* and *that* were further coded regarding their use in adverbial expressions as follows.

Table 8

Sample Coding of 'that' in Adverbial Expression L1 English and L2 English

Preceding text	POS	Ending text with reference	Source file
candles; others simply stand and weep. The demand for black clothes is so great	adv	that impromptu dyeing shops have sprung up, offering to turn brighter garments into someth	Thailand's succession - A royal mess.txt
bequests; set the rate high enough to raise significant sums, but not so high	adv	that it attracts massive avoidance. Third, with the fiscal headroom generated by higher in	A hated tax but a fair one - Inheritance tax.txt

Table 8 (cont'd)

electricity monopoly to buy coal only from black-owned firms; a process so mismanaged	adv	that it contributed to power cuts which knocked 1-2 percentage points off the national gro	South Africa's ruling party should dump Jacob Zuma - 783 reasons to go.txt
about food. In fact, the current situation about modern food is so ironic	adv	that among the questions about the subject, it might be natural to ask how people aided with	T4E3S (7).txt
becomes a stereotype. On the other hand, it can be argued that stereotypes are not all who succumb to the dark side of everything. But then again, maybe it's not	adv	that bad. Sure, there are stereotypes with negative connotations but this means that there that big a surprise. Why do people who start out as good people suddenly turn bad?	T2E3S (8).txt T5E3S (21).txt

In order to make a judgment for coding, it was sometimes necessary to consult the original text to gain more context. After a second or third read, the writers' intentions often became clearer, and the instances were appropriately coded. Because judgments regarding whether a pronoun is acting as a noun or an possessive, and whether a nominal pronoun is serving a subject, object or complement pronoun are rather straightforward, it was decided that a second coder was not necessary. In a few cases, after a second or third reading, some L2 writing remained too difficult or ungrammatical to be properly coded and were therefore coded 'z' and was excluded from analysis. Occurrences in either corpus that was part of quoted material and therefore did not reflect original writer content, were coded 'q' and likewise excluded from analysis.

Table 9

Sample Coding of 'z' in the L2 English Corpus

Preceding text	POS	Arg	Ending text with reference	Source file
he realised that that car is a possible killer for him because of algorithms ruling it. On the condition that being in	P	z	it into an oncoming traffic or a risky place like crossing a bridge, a mistaken or an	T1E3S (29).txt

Table 9 (cont'd)

nutritional malpractice needs to be addressed. There are many who commit to certain diets, in their pursuit of better health, or for their ideal body. This in	p	z	it of itself is not an issue, however the majority fail to visit an actual dietician, and get themselves an actual dietary plan made;	T4E3S (11).txt
started to arouse, comperatively to science fiction scenarios like The Terminator. Artificial Intelligence might seem helpful for humans and at	p	z	it the current state of A.I. it does not look like a huge menace to our society .Still, considering the potential "evolution"	T1E3S (10).txt
Thus, people should have information while they are consuming or they are feeding their children by these foods since	p	z	it they have influence on unhealthy future generations. At that position, home is a significant figure to	T4E3S (51).txt

Quantitative Analysis

Once the pronouns were sorted into useful categories and in order to gain a general understanding of differences in the two corpora, some initial calculations were made. Given the roughly equal size of the corpora, a cursory look at the magnitude of the raw numbers offered a quick impression of the differences. First, a search for terms in AntConc software provides a total count for the number of Occurrences for that term. AntConc also calculates the total size of the corpus (number of tokens or non-unique words). These numbers were manually entered into a spreadsheet to calculate the ratio of occurrences to total number of words in the corpus.

Table 10

Referential Pronouns: Raw Counts and Percentages

Pronoun	it		its		this		that	
	L1	L2	L1	L2	L1	L2	L1	L2
Occurrences	3,057	4,087	1,428	464	661	2,780	4,299	5,251
As % total corpus	1.0102%	1.0519%	0.4719%	0.1194%	0.2184%	0.7155%	1.4206%	1.3515%

Table 10 (cont'd)

Pronoun	these		those		they		them	
Corpus type	L1	L2	L1	L2	L1	L2	L1	L2
Occurrences	271	1,653	352	285	1,473	4,059	595	1,535
As % total corpus	0.0896%	0.4255%	0.1163%	0.0734%	0.4868%	1.0447%	0.1966%	0.3951%
Pronoun	his		hers		she		his/her	
Corpus type	L1	L2	L1	L2	L1	L2	L1	L2
Occurrences	1,034	1,141	2	0	323	173	0	41
As % total corpus	0.3417%	0.2937%	0.0007%	0.0000%	0.1067%	0.0445%	0.0000%	0.0106%
Pronoun	theirs		he		her		him	
Corpus type	L1	L2	L1	L2	L1	L2	L1	L2
Occurrences	6	3	1,121	638	369	218	202	173
As % total corpus	0.0020%	0.0008%	0.3704%	0.1642%	0.1219%	0.0561%	0.0668%	0.0445%

The higher frequency of the male and female personal pronouns *he*, *him*, *she*, and *her* in The Economist reflects its orientation towards news and personalities compared to the student L2 English corpus of argumentative essays. The only personal singular personal pronoun that occurred more frequently in the L2 English corpus was the ‘his/her’ construction, an attempt at gender neutrality still considered awkward by some and disfavored by The Economist Style Guide. It was therefore determined that such gender-specific pronouns would be excluded from the analysis of this study. In addition, due to its extremely low frequency, the pronoun *theirs* would also be excluded.

At first glance, some salient differences seemed to emerge. *They* and *them* seemed to occur at twice the rate in L2 English corpus as in the L1 English corpus; *its* occurs four times as frequently in the L1 English corpus as in the L2 English corpus; *that* occurs five times as frequently in the L2 as in the L1 English corpus. Further analysis based on syntactic function, part of speech (POS) and case (subject, object or complement) will reveal even deeper differences.

Table 11

Frequencies of 'this', 'that', 'these', and 'those' by Syntactic Function

	<u>this</u>		<u>that</u>		<u>these</u>		<u>those</u>	
	L1	L2	L1	L2	L1	L2	L1	L2
total	661	2780	4299	5251	271	1653	352	285
as % of total corpus	0,218%	0,716%	1,421%	1,352%	0,090%	0,425%	0,116%	0,073%
Nominative Expressions								
Nominative Pronoun	67%	33%	14%	7%	24%	9%	80%	23%
Argument								
subject	93%	75%	82%	56%	83%	66%	30%	31%
object	7%	22%	17%	36%	14%	27%	69%	66%
complement	0%	3%	1%	8%	3%	5%	2%	3%
unclear reference	0%	0%	0%	0%	0%	2%	0%	0%
Noun Clause								
	na	na	50%	60%	na	na	na	na
Generic Reference								
unqualified	na	na	na	na	na	na	1%	15%
qualified (non persona)	na	na	na	na	na	na	31%	20%
qualified reference to people	na	na	na	na	na	na	68%	65%
Adjective Expressions								
demonstrative adjective	32%	67%	5%	5%	76%	91%	20%	77%
adjective clause	na	na	29%	25%	na	na	na	na
Adverbial Expressions								
	1%	1%	2%	2%	na	na	na	na

For example, it appears that L1 English writers are many times less likely than L1 English writers to use *these* and *those* as nominative pronouns. At the same time, it appears that L2 English writers are three times as likely as L1 English writers to use the nominative pronoun *this* in an object position. These apparent differences would be more meaningful if they were tested for statistical significance.

To that end, it was necessary to test for statistical significance between the differences in the two corpora. The output from the AntConc program includes the source file where the reference occurred. Combined with the manual coding for that occurrence, a count was tabulated using Excel for each pronoun *by function, POS and case*. With this data, it was possible to calculate the means and standard deviations, and to perform an independent samples t-test for statistical significance between the means of the using IBM SPSS.

CHAPTER 4: DATA ANALYSIS

Introduction

This chapter gives the detailed analysis of the data from L1 and L2 corpora.

Data Analysis of Corpora Statistics

It

Table 12

Frequencies of 'it' by Syntactic Function

	L1 English	L2 English
Total Instances	3,057	4,087
Nominative Pronoun	3,057	4,087
<u>Argument</u>		
subject	2,351 (77%)	3,412 (83%)
object	699 (23%)	675 (17%)
complement	7 (0%)	0 (0%)
unclear reference	0 (0%)	47 (1%)
<u>Cleft or Fronted Structure</u>	529	1081
cleft subject	435 (83%)	1,037 (95%)
cleft object	94 (18%)	53 (5%)
<u>Non-referential</u>	55 (2%)	88 (2%)
<u>Time Expressions</u>	48 (2%)	22 (1%)

Occurrences of *it* were equally frequent in the corpora, a rate of 1.01% of the L1 English corpus and 1.05% of the L2 English corpus. L2 English writers were more likely to use the third-person impersonal pronoun in a subject position than L1 English writers.

It is significant to note that L2 English writers use *it* as part of fronted construction more frequently than L1 English writers do (26% of occurrences compared to 17%). When using a cleft structure, the student writers overwhelmingly

favor using the cleft in a subject position, while L2 English writers are much more inclined than L2 English writers to put it in an object position. By way of example, an occurrence of *it* such as the following is classified as a cleft subject: “It is essential that French voters have a decent alternative to Ms Le Pen” (The French presidential election - Europe's biggest populist danger); while the following instance is classified as a cleft object: “The disabled, the old and the young will find it easier to go where they want” (Personal Transportation – Uberworld).

Table 13

Group Statistics for 'it' by Syntactic Function

It	Corpus	N	Mean	Std. Deviation	Std. Error Mean
Subject	L1 English	380	6.19	3.165	0.162
	L2 English	372	9.17	6.168	0.32
Complement	L1 English	380	0	0.051	0.003
	L2 English	372	0	0	0
Object	L1 English	380	1.83	1.575	0.081
	L2 English	372	1.81	2.268	0.118
Cleft/Fronted	L1 English	380	1.39	1.296	0.066
	L2 English	372	2.77	2.706	0.14
Time Expression	L1 English	380	0.13	0.377	0.019
	L2 English	372	0.06	0.247	0.013
Idiomatic	L1 English	380	0.12	0.352	0.018
	L2 English	372	0.12	0.428	0.022
Non-referential	L1 English	380	0.03	0.176	0.009
	L2 English	372	0.09	0.332	0.017

Table 14

T-test Statistics for 'it' by Syntactic Function

It	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
Subject	125.023	.000	-8.368	750	.000	-2.983	.356
Complement	3.936	.048	.989	750	.323	.003	.003
Object	26.491	.000	.139	750	.890	.020	.142
Cleft /	144.964	.000	-8.960	750	.000	-1.382	.154

Fronted							
Time exp	34.150	.000	2.881	750	.004	.067	.023

Table 14 (cont'd)

Idiomatic	.029	.865	.007	750	.994	.000	.029
Non-referent	47.134	.000	-3.368	750	.001	-.065	.019

There is a statistically significant difference in use of *it* in subject position between L1 (M=6.19, SD=3.17) and L2 (M=9.17, SD=6.17) writing ($t(750.000)=-8.368, p<.05$).

There is no statistically significant difference in use of *it* in complement position between L1 (M=0.00, SD=0.51) and L2 (M=0.00, SD=0.00) writing ($t(750.000)=.989, p>.05$).

There is no statistically significant difference in use of *it* in object position between L1 (M=1.83, SD=1.58) and L2 (M=1.81, SD=2.27) writing ($t(750.000)=0.139, p>.05$).

There is a statistically significant difference in use of *it* as a cleft between L1 (M=1.39, SD=1.30) and L2 (M=2.77, SD=2.71) writing ($t(750.000)=-8.960, p<.05$).

There is a statistically significant difference in use of *it* as a time expression between L1 (M=0.13, SD=0.38) and L2 (M=0.06, SD=0.25) writing ($t(750.000)=2.881, p<.05$).

There is no statistically significant difference in use of *it* as idiomatic usage between L1 (M=0.12, SD=0.35) and L2 (M=0.12, SD=0.43) writing ($t(750.000)=0.007, p>.05$). Both NS and NNS writers used a variety time and expressions (It was the late 1980s; It was too late; It takes too long) in equal proportion.

There is a statistically significant difference in use of *it* as a non-referential expression between L1 (M=0.03, SD=0.18) and L2 (M=0.09, SD=0.33) writing ($t(750.000)=-3.368, p<.05$). The use of these expressions was distributed throughout

sections, suggesting a universal ability to effectively construct time expressions using the non-referential *it*.

Table 15

L2 Concordance Analysis for 'it' in time expressions

Preceding text	POS	Arg.	Type	Ending text with reference	Source file
with less doubts related to ingredients and preparing paths of food. He argues that,	p	s	time	it is not too much time to have an healthier and longer life. People do not cook at their home	T4E3S (32).txt
The struggle should be done to prevent the end of human world before	p	s	time	it is too late.	T1E3S (19).txt
in this developing technological era, Macintosh introduced Siri (SRI International) to world and	p	s	time	it was the first time the artificial intelligence becomes a part of daily life. Artificial intelligence (AI)]	T1E3S (40).txt
stereotypes are trying to change society in the way that they want. Otherwise,	p	s	time	it will be too late when we realize that we feel like we are not accepted or we do not have a room	T2E3S (51).txt
The worst part of this is that man doesn't realize that he is in trouble until	p	s	time	it's too late. Everyone thinks that it is only a small thing it can't hurt me or anyone else.	T5E3S (40).txt

Table 16

L1 Concordance analysis for 'it' in time expressions

Preceding text	POS	Arg.	Type	Ending text with reference	Source file
And the man in the Oval Office is making a bad situation worse. JULY 4th ought to bring Americans together.	p	s	time	It is a day to celebrate how 13 young colonies united against British rule to begin	Donald Trump's Washington is paralysed - American politics.txt
	p	s	time	IT HAS been many years since France last had a revolution, or even a serious	The vote that could wreck the European Union - France's next revolution.txt
the first immunotherapy treatment in effect disables the brakes, enabling white blood cells to attack the tumours.	p	s	time	It is early days, but in a small subset of patients this mechanism has produced	Closing in on cancer - Health care.txt
Richard Burr, Ben Sasse and John McCain, are troubled by what the removal of Mr Comey portends.	p	s	time	It is high time for them and others to put their country before their party.	After Comey's dismissal, it's time for a commission - Russia and the Trump campaign.txt

Table 16 (cont'd)

effects of poison gases used on them by Iraq in the war of 1980-88. The longer Russia and Iran keep Mr Assad in power, the more they will share in his guilt.	p	s	time	It is time for them to ditch their toxic ally. This article appeared in the	Why Russia and Iran should ditch Bashar al-Assad - A poisonous client.txt
---	---	---	------	---	---

One-percent of the *it* referents from the L2 English corpus were coded as 'unclear'. These did not occur in the L1 English corpus. To follow are some examples.

Table 17

Coding 'it' for Unclear References

Preceding text	POS	Arg.	Type	Ending text with reference	Source file
s which are required to be neutral can cause serious problems. At this point governments' intervention is right decision to save social peace. This intervention is not about biases or discrimination,	p	s	unclear reference	it is about saving equality and secularity. However, these interventions should not be extreme level and should not limit people's freedom. Governments responsible for citizens welfare, they should no	T6E3S (43).txt
preferred women who would laugh at their jokes to those who made jokes. Women, however, preferred partners who were funny..."(2). In other words, as society thinks women likes men who are funny,	p	s	unclear reference	it is above the rumour and that is more deeper. However, the person who makes the jokes must be men. In addition, in The Humor Gap, according to Nicholson, it is explained that "	T7E3S (1).txt
like obesity. However, people are getting more knowledgeable day by day. That can be a break point for healthy eating habit for the future generations. These problems are not children's fault,	p	s	unclear reference	it is all about parent's actions but there is hope for healthy future generation. Eat healthy, stay healthy.	T4E3S (46).txt
. Another example, children can be jealous and they can harm their friends toys or bodies, but, it does not make them evil because they are not able to think what will be happened.	p	s	unclear reference	It is another situation to act like a demon because it is not possible to predict what is bad or good and it can creates bad situations, so these actions can be bad. Conditions	T5E3S (26).txt
is evil because the situation, the factors made them evil. Humanity is like a clean sheet of paper. If It is in a dirty place, it will get dirty in the end. If	p	s	unclear reference	it is in an immoral environment, the ultimate evil will be born from the ashes of it.	T5E3S (5).txt
cial for people. Religions occurred thousands of years ago to provide order to people, but especially to people in that age. Rules or implementations that religions may be beneficial in the past, but	p	s	unclear reference	it is likely them to contradict today's understanding of freedom or basic human rights. Death penalties, for example, were the most common punishment for many crimes in the past, whereas it is	T6E3S (14).txt

Table 17 (cont'd)

he society and people or cause people to lose their rights .government should intervene in these religious differences to some extent. This intervention shouldn't be more than necessity because if	p	s	unclear reference	it is more this cause people to lose their religious rights and freedoms and exclusion of people who have different religious beliefs in society. Everyone has freedom of religion and belief but i	T6E3S (7).txt
---	---	---	----------------------	---	------------------

This apparent difficulty in effective referencing with *it* and subsequent loss of cohesion (and coherence) is consistent with the findings of and Kim (2012) who reported Korean EFL writers' tendency to use *it* to refer to long sentences.

Its

Table 18

Concordance Analysis for 'its'

	L1 English	L2 English
Total Instances	1,428	464
as % of total corpus	0.472%	0.119%

Use of the possessive pronoun *its* is used much less frequently and represents a much lower percentage of the overall corpus (0,5% of L1 English corpus vs. 0,12% of L2 English corpus).

Table 19

Group Statistics for 'its'

	Corpus	N	Mean	Std. Deviation	Std. Error Mean
its	L1 English	380	3.76	3.158	0.162
	L2 English	372	1.24	1.971	0.102

Table 20

T-test findings for 'its'

	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
its	66.153	.000	13.087	750	.000	2.519	.192

There is a statistically significant difference in use of *its* between L1 (M=3.76, SD=3.16) and L2 (M=1.24, SD=1.49) writing ($t(750.000)=13.087, p<.05$).

This

Table 21

Frequencies of 'this' by Syntactic Function

	L1 English	L2 English
Total Instances	661 (0.218%)	2,780 (0.716%)
Nominative Pronoun	444 (67%)	912 (33%)
<u>Argument</u>		
subject	414 (93%)	684 (75%)
object	30 (7%)	202 (22%)
complement	0 (0%)	26 (3%)
unclear reference	0 (0%)	2 (0%)
<u>Adjective Expressions</u>	529	1081
demonstrative adjective	212 (32%)	1,862 (67%)
<u>Adverbial Expressions</u>	5 (1%)	6 (1%)

Turkish L2 English writers depend heavily on the pronoun *this* to achieve cohesion in their writing. The pronoun represents 0.2% of the L1 English corpus but 0.7% of the L2 English corpus. L1 English writers have a strong tendency to use *this* as a subject case pronoun; L2 English writers have the same tendency, though not as pronounced – 93% vs. 75% of occurrences. Similarly, L2 English writers use *this* more frequently as an object case pronoun (22% of all uses as a nominative pronoun) compared to L1 English writers (7% of all uses as a nominative pronoun). The proportion of uses as a nominative pronoun compared to demonstrative adjective are

perfectly reversed between L1 and L2 English writers: 67% of occurrences of *this* are nominative pronouns in L2 writing, while 67% of occurrences in L1 writing are as demonstrative adjectives. L1 and L2 English writers appear to use *this* in adverbial expressions in equal proportions (1% of all expressions of *this*).

Table 22

Group Statistics for 'this' by Syntactic Function

It	Corpus	N	Mean	Std. Deviation	Std. Error Mean
This Demo. Pronoun	L1 English	380	1.16	1.215	0.062
	L2 English	372	2.45	2.388	0.124
This Demo. Pronoun Subject	L1 English	380	0.14	0.379	0.019
	L2 English	372	0.25	0.627	0.032
This Demo. Pronoun Compl.	L1 English	380	0.01	0.072	0.004
	L2 English	372	0.02	0.136	0.007
This Demo. Pronoun Object	L1 English	380	0.02	0.152	0.008
	L2 English	372	0.1	0.403	0.021
This Demo. Adjective	L1 English	380	1.75	1.5	0.077
	L2 English	372	4.99	4.118	0.213
This Adverbial Expression	L1 English	380	0.01	0.114	0.006
	L2 English	372	0.02	0.126	0.007

Table 23

T-test Findings for 'this' by Syntactic Function

This	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
Demonstrative Pronoun	101.401	.000	-9.372	750	.000	-1.291	.138
Demonstrative Adjective	150.183	.000	-14.415	750	.000	-3.245	.225
Adverbial	.460	.498	-.339	750	.735	-.003	.009
Demonstrative Pronoun Subject	31.496	.000	-2.864	750	.004	-.108	.038
Demonstrative Pronoun Complement	11.876	.001	-1.710	750	.088	-.014	.008
Demonstrative Pronoun Object	51.882	.000	-3.549	750	.000	-.078	.022

There is a statistically significant difference in use of *this* as a demonstrative pronoun between L1 (M=1.16, SD=1.22) and L2 (M=2.45, SD=2.39) writing ($t(750.000) = -9.372, p < .05$).

There is a statistically significant difference in use of *this* as a demonstrative adjective between L1 (M=1.75, SD=1.50) and L2 (M=4.99, SD=4.19) writing ($t(750.000) = -14.415, p < .05$).

There is no statistically significant difference in use of *this* as an adverbial expression between L1 (M=0.01, SD=0.11) and L2 (M=0.02, SD=0.17) writing ($t(750.000) = -.339, p > .05$).

There is a statistically significant difference in use of *this* as a demonstrative pronoun in the subject position between L1 (M=0.14, SD=0.379) and L2 (M=0.25, SD=0.627) writing ($t(750.000) = -2.864, p > .05$).

There is no statistically significant difference in use of *this* demonstrative pronoun in the complement position between L1 (M=0.01, SD=0.072) and L2 (M=0.02, SD=0.136) writing ($t(750.000) = -1.710, p > .05$).

There is a statistically significant difference in use of *this* as a demonstrative pronoun in the object position between L1 (M=0.02, SD=0.152) and L2 (M=0.10, SD=0.403) writing ($t(750.000) = -3.549, p > .05$).

Because of the L1 English corpus's focus on news events, there was a disproportionate number of time reference collocations using *this* (e.g. 'this week', 'this month,' and 'this year'). Such references, which numbered 263 of demonstrative adjectives or more than 50% of all such adjective expressions in the L1 English corpus, were unlikely to occur in an academic essay and were therefore excluded from the analysis. Similarly, 34 occurrences self-referential expressions of a strictly editorial style ('this newspaper') were also excluded from the analysis.

That

The pronoun *that* generated the greatest number of hits in both the L1 and L2 corpora, a little over 4,000 and well over 5,000 respectively. This may be attributed to the fact that the word has multiple functions in English syntax. It took painstaking attention to sort the instances of *that* as a demonstrative pronoun, demonstrative adjective (also called a demonstrative determiner), or whether it was used to introduce a noun clause, a relative clause, or an adverb clause.

Table 24

Frequencies of 'that' by Syntactic Function

	L1 English	L2 English
<u>Nominative Pronoun</u>	598 (14%)	600 (7%)
<u>Argument</u>		
subject	489 (82%)	432 (56%)
object	104 (17%)	139 (36%)
complement	5 (1%)	29 (8%)
idiomatic/non referential	1 (0%)	1 (0%)
<u>Noun Clause</u>	2,155 (50%)	3,158 (60%)
<u>Demonstrative Adjective</u>	212 (5%)	273 (5%)
<u>Relative Clause</u>	1,252 (29%)	1,309 (25%)
<u>Adverbial Expressions</u>	5 (2%)	6 (2%)

The proportion of occurrences that introduced noun clauses, adjective clauses and adverbial expressions were roughly equivalent.

Table 25

Group Statistics for 'that' by Syntactic Function

	Corpus	N	Mean	Std. Deviation	Std. Error Mean
Adjective Clause	L1 English	380	3.29	2.074	0.106
	L2 English	372	3.52	3.64	0.189

Table 25 (cont'd)

Adverbial Expression	L1 English	380	0.22	0.46	0.024
	L2 English	372	0.34	0.663	0.034
Demonstrative Adjective	L1 English	380	0.56	0.765	0.039
	L2 English	372	8.49	5.376	0.279
Demonstrative Pronoun	L1 English	380	1.57	1.263	0.065
	L2 English	372	0.73	1.288	0.067
Noun Clause	L1 English	380	5.67	3.156	0.162
	L2 English	372	1.03	1.458	0.076
Demonstrative Pronoun Subject	L1 English	380	1.29	1.164	0.06
	L2 English	372	0.58	1.203	0.062
Demonstrative Pronoun Compl.	L1 English	380	0.01	0.114	0.006
	L2 English	372	0.08	0.318	0.017
Demonstrative Pronoun Object	L1 English	380	0.27	0.528	0.027
	L2 English	372	0.37	0.725	0.038

Table 26

T-test findings for 'that' by Syntactic Function

That	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
Adjective clause	60.661	.000	-1.040	750	.299	-.224	.215
Adverbial expressions	36.786	.000	-2.957	750	.003	-.123	.042
Demonstrative adjective	428.348	.000	-28.478	750	.000	-7.934	.279
Demonstrative Pronoun	8.899	.003	9.057	750	.000	.843	.093
Noun clause	146.801	.000	25.772	750	.000	4.636	.180
Demonstrative Pronoun Subject	6.453	.011	8.182	750	.000	.706	.086
Demonstrative Pronoun Complement	64.561	.000	-3.886	750	.000	-.067	.017
Demonstrative Pronoun Object	19.677	.000	-2.165	750	.031	-.100	.046

There is no statistically significant difference in use of *that* to introduce an adjective clause between L1 (M=3.29, SD=2.07) and L2 (M=3.52, SD=3.64) writing ($t(586.426) = -1.040, p > .05$)

There is a statistically significant difference in use of *that* as an adverbial expression between L1 (M=0.22, SD=0.46) and L2 (M=0.34, SD=0.66) writing ($t(750.000) = -2.975, p < .05$).

There is a statistically significant difference in use of *that* as a demonstrative adjective between L1 (M=0.56, SD=0.77) and L2 (M=8.49, SD=5.38) writing ($t(750.000) = -28.478, p < .05$).

There is a statistically significant difference in use of *that* as a demonstrative pronoun in all cases between L1 (M=1.57, SD=1.26) and L2 (M=0.73, SD=1.29) writing ($t(750.000) = 9.057, p < .05$).

There is a statistically significant difference in use of *that* to introduce a noun clause between L1 (M=5.67, SD=3.16) and L2 (M=1.03, SD=1.46) writing ($t(750.000) = 25.772, p < .05$).

There is a statistically significant difference in use of *that* as a demonstrative pronoun in subject position between L1 (M=1.29, SD=1.16) and L2 (M=0.58, SD=1.20) writing ($t(747.843) = 8.182, p < .05$).

There is a statistically significant difference in use of *that* as a demonstrative pronoun in complement position between L1 (M=0.01, SD=0.11) and L2 (M=0.08, SD=0.32) writing ($t(750.000) = -3.886, p < .05$).

There is a statistically significant difference in use of *that* as a demonstrative pronoun in object position between L1 (M=0.27, SD=0.53) and L2 (M=0.37, SD=0.73) writing ($t(750.000) = -2.165, p < .05$).

These

Table 27

Frequencies of 'these' by Syntactic Function

	L1 English	L2 English
Total instances	271 (0,090%)	1,653 (0,425%)
Nominative Pronoun	65 (24%)	141 (9%)
<u>Argument</u>		
subject	54 (83%)	93 (66%)
object	9 (14%)	38 (27%)
complement	2 (3%)	7 (5%)
idiomatic/non referential	0 (0%)	3 (2%)
<u>Adjective Expressions</u>		
demonstrative adjective	206 (76%)	1,512 (91%)

As is the case with its singular analog *this*, *these* functions as either a nominal or adjective pronoun describing things in close proximity. L1 writing contains over four times as many instances of *these* as L2 writing. Of those occurrences, *these* was used 91% of the time as an adjective among NNS compared to 76% for NS writers.

As a nominal pronoun, *these* was used only 9% of the time by NNS and 24% by NS. Student writers used *these* as a subject case pronoun 66% of the time compared to L1 English corpus, where writers used *these* as a subject case pronoun in 83% of the cases where it was a nominal pronoun.

Table 28

Group Statistics for 'these' by Syntactic Function

	Corpus	N	Mean	Std. Deviation	Std. Error Mean
Demonstrative Pronoun	L1 English	380	0.17	0.429	0.022
	L2 English	372	0.37	0.816	0.042
Demonstrative Pronoun Subject	L1 English	380	0.14	0.379	0.019
	L2 English	372	0.25	0.627	0.032

Table 28 (cont'd)

Demonstrative Pronoun Complement	L1 English	380	0.01	0.072	0.004
	L2 English	372	0.02	0.136	0.007
Demonstrative Pronoun Object	L1 English	380	0.02	0.152	0.008
	L2 English	372	0.1	0.403	0.021
Demonstrative Adjective	L1 English	380	0.54	0.773	0.04
	L2 English	372	4.06	3.506	0.182

Table 29

T-test findings for 'these' by Syntactic Function

These	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
Demonstrative Pronoun	59.926	.000	-4.218	750	.000	-.200	.047
Demonstrative Pronoun Subject	31.496	.000	-2.864	750	.004	-.108	.038
Demonstrative Pronoun Complement	11.876	.001	-1.710	750	.088	-.014	.008
Demonstrative Pronoun Object	51.882	.000	-3.549	750	.000	-.078	.022
Demonstrative Adjective	338.031	.000	-19.130	750	.000	-3.525	.184

There is a statistically significant difference in use of *these* as a demonstrative pronoun all cases between L1 (M=0.17, SD=0.43) and L2 (M=0.37, SD=0.82) writing ($t(750.000) = -4.218, p < .05$).

There is a statistically significant difference in use of *these* as a demonstrative pronoun in subject position between L1 (M=0.14, SD=0.38) and L2 (M=0.25, SD=0.63) writing ($t(750.000) = -2.864, p < .05$).

There is no statistically significant difference in use of *these* as a demonstrative pronoun in complement position between L1 (M=0.01, SD=0.07) and L2 (M=0.02, SD=0.14) writing ($t(750.000) = -1.710, p > .05$).

There is a statistically significant difference in use of *these* as a demonstrative pronoun in object position between L1 (M=0.02, SD=0.15) and L2 (M=0.10, SD=0.40) writing ($t(750.000) = -3.549, p < .05$).

There is a statistically significant difference in use of *these* as a demonstrative adjective between L1 (M=0.54, SD=0.773) and L2 (M=4.06, SD=3.51) writing ($t(750.000) = -19.130, p < .05$).

Those

As described in the Methodology section, *those* plays a number of roles in English syntax and a number of special roles in reference and cohesion in English language discourse. In its role as a demonstrative pronoun (dp), it is used to refer to previously mentioned objects or persons, and often with a qualifying adjective clause or phrase attached. Such uses were coded to distinguish them from basic, non-qualified expressions.

Table 30

Frequencies of 'those' by Syntactic Function

	L1 English	L2 English
Total instances	352 (0.116%)	285 (0.073%)
Nominative Pronoun	281 (80%)	65 (23%)
<u>Argument</u>		
subject	83 (30%)	20 (31%)
object	193 (69%)	43 (66%)
complement	5 (2%)	2 (3%)
non-qualified	0 (0%)	10 (15%)
qualified non-personal	86 (31%)	13 (20%)
qualified personal	195 (69%)	42 (65%)
<u>Adjective Expressions</u>		
demonstrative adjective	71 (20%)	220 (77%)

Table 31

Group Statistics for 'those' by Syntactic Function

	Corpus	N	Mean	Std. Deviation	Std. Error Mean
Demonstrative Pronoun	L1 English	380	0.74	0.995	0.051
	L2 English	372	0.17	0.537	0.028
Demonstrative Pronoun Subject	L1 English	380	0.22	0.49	0.025
	L2 English	372	0.05	0.285	0.015
Demonstrative Pronoun Complement	L1 English	380	0.01	0.114	0.006
	L2 English	372	0.01	0.073	0.004
Demonstrative Pronoun Object	L1 English	380	0.51	0.791	0.041
	L2 English	372	0.12	0.367	0.019
Demonstrative Pronoun Qualified Personal	L1 English	380	0.53	0.92	0.047
	L2 English	372	0.11	0.444	0.023
Demonstrative Pronoun Qualified Non-Personal	L1 English	380	0.21	0.459	0.024
	L2 English	372	0.03	0.185	0.01
Demonstrative Non-qualified	L1 English	380	0	0	0
	L2 English	372	0.03	0.17	0.009
Demonstrative Adjective	L1 English	380	0.19	0.453	0.023
	L2 English	372	0.59	1.288	0.067

Table 32

T-test Findings for 'those' by Syntactic Function

Those	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
Demonstrative Adjective	86.397	.000	-5.768	750	.000	-.405	.070
Demonstrative Pronoun	125.121	.000	9.703	750	.000	.567	.058
Demonstrative Pronoun Subject	134.173	.000	5.715	750	.000	.167	.029
Demonstrative Pronoun Complement	4.965	.026	1.110	750	.267	.008	.007
Demonstrative Pronoun Object	215.490	.000	8.695	750	.000	.392	.045
Qual. Ref. to Persons	175.911	.000	7.820	750	.000	.413	.053

Table 32 (cont'd)

Qual. Ref to Non-persons	248.130	.000	7.171	750	.000	.184	.026
Non-qualified Reference	49.142	.000	-3.398	750	.001	-.030	.009

Those as a Demonstrative Adjective

There is a statistically significant difference in use of *those* as a demonstrative adjective between L1 (M=0.19, SD=0.453) and L2 (M=0.59, SD=1.29) writing ($t(750.000) = -5.768, p < .05$). L2 English writers are much more likely than L1 English writers to use *those* as a demonstrative adjective before a noun to refer to something previously stated, often reflecting the topic of the essay: “those robots”, “those stereotypes”, or something more generic: “those people”, “those situations”. It may be worth noting here that L2 English writers are many, many times more likely to use ‘people’ as a generic reference for the third person as well. In fact, ‘people’ is the 9th most frequently occurring token in the L2 English corpus (5080 occurrences), compared to 53rd in the L1 English corpus (607 occurrences). L2 English writers rely too heavily on this generic “catch-all” reference instead of a more precise enumerative term. This may be related to limited vocabulary, or transfer of an L1 writing style that is more tolerant of this generic term.

Those as a Demonstrative Pronoun

There is a statistically significant difference in use of *those* as a demonstrative pronoun all cases between L1 (M=0.74, SD=1.00) and L2 (M=0.17, SD=0.54) writing ($t(750.000) = 9.703, p < .05$).

There is a statistically significant difference in use of *those* as a demonstrative pronoun in subject position between L1 (M=0.22, SD=0.50) and L2 (M=0.05, SD=0.29) writing ($t(750.000) = 5.715, p < .05$).

There is no statistically significant difference in use of *those* as a demonstrative pronoun in complement position between L1 (M=0.01, SD=0.11) and L2 (M=0.01, SD=0.07) writing ($t(750.000)= 1.110, p>.05$).

There is a statistically significant difference in use of *those* as a demonstrative pronoun in object position between L1 (M=0.51, SD=0.79) and L2 (M=0.12, SD=0.37) writing ($t(750.000)= 8.695, p<.05$).

There is a statistically significant difference in use of *those* as a qualified reference to persons between L1 (M=0.53, SD=0.92) and L2 (M=0.11, SD=0.44) writing ($t(750.000)= 7.820, p<.05$).

There is a statistically significant difference in use of *those* as a qualified reference to non-persons between L1 (M=0.21, SD=0.46) and L2 (M=0.03, SD=0.19) writing ($t(750.000)= 7.171, p<.05$).

The pronoun *those* is more widely used in the L1 English corpus, comprising 0,11% of all words compared to 0,07 in the L2 English corpus. In the L1 English corpus, 80% of the occurrences were nominal pronouns compared to 23% in the L2 English corpus. Despite this disparity, the distribution of subject, object and complements were rather even.

There is a statistically significant difference in use of *those* as non-qualified demonstrative pronoun L1 (M=0.00, SD=0.00) and L2 (M=0.03, SD=0.17) writing ($t(750.000)= -3.398, p<.05$).

Non-qualified nominal expressions occur exclusively in the L2 English corpus, never in the L1 English corpus. For example:

Table 33

L2 English Writers' use of 'those' as a Non-qualified Nominative Pronoun

however, when people behave towards people as group or "herd" in order to they share responsibility, they feel less responsible towards their behavior which is evil. When	<u>those</u> are doing bad things as a group, they assume that this behavior moral and does not harm others. In order to their psychology is affected negatively, they do not consider whether they are
sources of inconsistent information about food and diets, which only cooperates to make this situation even worse. Although many options to provide healthier options of food are being developed,	<u>those</u> are not available to a large share of the world population yet. In short, a revolution in modern eating habits and trends is urgently necessary, in order to avoid creating future generations
The disadvantages should not lead people to stop using it. Even though	<u>those</u> are real facts, the advantages on the other hand makes social media the platform of opportunities.
These private contents can even affect their future; if	<u>those</u> come out one day, their school or job applications can even be canceled,
Artificial intelligence can awaken like a candy. However , AI inevitably poses lots of risk and danger .AI cannot be used regardless of risks and threats. If	<u>those</u> compare benefits and dangers of artificial intelligence by taking into account, they can easily realize that dangers of artificial intelligence outweighs its benefits.
In conclusion, nobody should stop using social media, because while it may have risks for adolescents,	<u>those</u> could be solved with proper education to both parents and adolescents,
Maybe, they will not eat high-calorie snacks if they know how Maybe, they will not eat high-calorie snacks if they know how	<u>those</u> have the energy equivalent to a meal. All that being said, providing information is a simple way to solve the health problems of future generations.

Even with extended text before the reference, the usage seems stilted. It seems much more natural to use a qualifying expression, whether for persons or objects, to convey the precise meaning of the pronoun, often as a subgroup of a previously mentioned one. These qualifiers are either adjective clauses or a reduction thereof.

Table 34

L1 English Writers' Use of 'those' as a Qualified Nominative Expression

government lets them throw their names in. Companies can trawl that pool for workers;	those they sponsor will shoot up the rankings and get in more quickly. This is
---	--

Table 34 (cont'd)

atomic clocks by incorporating entanglement, for example, makes them more accurate than	those used today in satellite positioning. That could improve navigational precision by or
whether they are new or established, prevents the research that could distinguish between	those which are more and less harmful. It also leads to topsy-turvy outcomes.
contributes to the problem. Even eliminating all subsidies, not just	those which finance dodgy activities, will not on its own spare the oceanic commons from
of more than 150 officials deemed connected to the drug trade, at least two of	those whom he fingered were already dead. It would be comical were the consequences not
Widely accessible sites could be more tightly regulated than	those with a restricted audience. The drawback is that this turns online firms, especially

In these examples of L1 English writing, the referent attached to *those* is in very close proximity, making resolution much easier, while the qualifying information narrows the meaning. This technique did occur in the L1 writer corpus, but less frequently and less effectively. Fifteen percent of the nominal expressions *those* in L2 English corpus are non-qualified. Qualified references to persons were roughly equal (68% and 65% of nominal expressions for L1 and L2 respectively) while qualified references to non-persons represented 31% of pronominal references for L1 and 20% for L2.

They

On a proportional basis, the third person plural subject case pronoun was used twice as frequently by L2 English writers than by L1 English writers. There is a statistically significant difference in use of *they* L1 ($M=3.88$, $SD=3.14$) and L2 ($M=10.91$, $SD=8.89$) writing ($t(460.232)=-14.404$, $p<.05$).

Table 35

Frequencies of 'they'

	L1 English	L2 English
Total Instances	1,473 (0.487%)	4,059 (1.045%)
<u>Nominative Expressions</u>	1,473	4,059
Argument		
subject	1,473	4,059

Table 36

Group Statistics for 'they'

	Corpus	N	Mean	Std. Deviation	Std. Error Mean
they	L1 English	380	3.88	3.136	0.161
	L2 English	372	10.91	8.89	0.461

Table 37

T-test Findings for 'they'

	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
they	184.005	.000	-14.524	750	.000	(7.032)	.484

There is a statistically significant difference in use of *they* L1 (M=3.88, SD=3.138) and L2 (M=10.91, SD=8.89) writing ($t(750.000) = -14.524, p < .05$). There are sizeable and significant differences in the occurrence of *they* between the two corpora. Without a deeper level of discourse analysis, it is difficult to explain this difference. There may be topical effect on the genres that influence the use of the pronoun, or the difference may be related to L2 interference.

Them

Table 38

Frequencies of 'them'

	L1 English	L2 English
Total Instances	595 (0.197%)	1,535 (0.395%)
<u>Nominative Expressions</u>	1,473	4,059
Argument		
object	595	1,535

Table 39

Group Statistics for 'them'

	Corpus	N	Mean	Std. Deviation	Std. Error Mean
them	L1 English	380	1.57	1.468	0.075
	L2 English	372	3.63	3.164	0.164

Table 40

T-test Findings for 'them'

	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
them	114.987	.000	-11.478	750	.000	(2.058)	.179

There is a statistically significant difference in use of *them* L1 (M=1.57, SD=1.47) and L2 (M=3.63, SD=3.16) writing ($t(750.000) = -11.478, p < .05$). On a proportional basis, the third person plural object case pronoun, like the subject case pronoun, was used much more frequently, the reasons for which may be understood with further analysis.

Their

Table 41

Frequencies of 'their'

	L1 English	L2 English
Total Instances	1170 (0.387%)	3285 (0.846%)
<u>Adjective Expressions</u>	1,473	4,059
Argument		
possessive adjective	1,170	3,285

Table 42

Group Statistics for 'their'

	Corpus	N	Mean	Std. Deviation	Std. Error Mean
their	L1 English	380	3.08	2.61	0.134
	L2 English	372	8.83	5.982	0.31

Table 43

T-test Findings for 'their'

	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
their	200.124	.000	-17.151	750	.000	-5.752	.335

There is a statistically significant difference in use of *their* L1 (M=3.08, SD=2.61) and L2 (M=8.83, SD=5.98) writing ($t(750.000) = -17.151, p < .05$). As with the other third-person plural pronouns *they* and *them*, this higher occurrence – more than double the rate of the L1 writer corpus – may reflect the influence of L2, or an effect of the genre or topic. Further analysis is required to know.

Theirs

Table 44

Frequencies of 'theirs'

	L1 English	L2 English
Total instances	6 (0.002%)	3 (0.001%)
Nominative Pronoun	6 (33%)	3 (67%)
<u>Argument</u>		
subject	2 (50%)	2 (33%)
object	3 (17%)	1 (0%)
complement	1	0

Table 45

Group Statistics for 'theirs'

	Corpus	N	Mean	Std. Deviation	Std. Error Mean
theirs	L1 English	380	0.02	0.125	0.006
	L2 English	372	0.01	0.09	0.005

Table 46

T-test Findings for 'theirs'

	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
theirs	3.809	.051	.973	750	.331	.008	.008

There is no statistically significant difference in use of *theirs* between L1 (M=0.02, SD=0.13) and L2 (M=0.01, SD=0.09) writing ($t(750.000)=0.973$, $p>.05$). Frequency of use was too small to draw any meaningful conclusions.

Third Person Singular Personal Pronouns**He.**

Table 47

Frequencies of 'he'

	L1 English	L2 English
Total instances	1,121 (0.370%)	638 (0.164%)
Nominative Expressions	1,121 (33%)	638 (67%)
<u>Argument</u>		
subject	1,121	638

Table 48

Group Statistics for 'he'

	Corpus	N	Mean	Std. Deviation	Std. Error Mean
Table 48 (cont'd)					
he	L1 English	380	2.95	4.333	0.222
	L2 English	372	1.71	3.072	0.159

Table 49

T-test Findings for 'he'

	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
he	28.384	.000	4.520	750	.000	1.240	.274

The third person masculine subject case singular pronoun was twice as frequent in the L1 English corpus as the L2 English corpus.

There is a statistically significant difference in use of *he* between L1 (M=2.95, SD=4.33) and L2 (M=1.71, SD=3.07) writing ($t(750.000) = 4.520, p < .05$).

Her.

Table 50

Frequencies of 'her' by Syntactic Function

	L1 English	L2 English
Total Instances	369 (0,122%)	218 (0,056%)
Nominative Pronoun Argument	62 (17%)	47 (22%)
subject	0 (0%)	0 (0%)
object	62 (100%)	47 (100%)
complement	0 (0%)	0 (0%)
<u>Adjective Expressions</u>		
possessive adjective	308 (83%)	171 (78%)

Table 51

Group Statistics for 'her' by Syntactic Function

Her	Corpus	N	Mean	Std. Deviation	Std. Error Mean
Personal Pronoun	L1 English	380	0.16	0.633	0.032
	L2 English	372	0.13	0.392	0.02
Possessive Adjective	L1 English	380	0.81	2.516	0.129
	L2 English	372	0.46	1.147	0.059

Table 52

T-test Findings for 'her'

Her	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
Personal Pronoun	4.736	.030	.956	750	.339	.037	.038
Possessive Adjective	23.391	.000	2.452	750	.014	.351	.143

The pronoun *her* serves both a nominative and possessive function. Though twice as frequent in L1 English corpus, the distribution of nominative and possessive functions are equivalent.

There is no statistically significant difference in use of *her* as an object case pronoun between L1 (M=0.16, SD=0.63) and L2 (M=0.13, SD=0.39) writing ($t(750.000)=.956, p>.05$).

There is a statistically significant difference in use of *her* as possessive adjective between L1 (M=0.81, SD=2.52) and L2 (M=0.46, SD=1.15) writing ($t(750.000)= 2.452, p<.05$).

Him.

Table 53

Frequencies of 'him'

	L1 English	L2 English
Total Instances	202 (0.067%)	173 (0.045%)
<u>Nominative Expressions</u>		
Argument		
	object	
	202	173

Table 54

Group Statistics for 'him'

him	Corpus	N	Mean	Std. Deviation	Std. Error Mean
him	L1 English	380	0.53	1.117	0.057
	L2 English	372	0.47	1.162	0.06

Table 55

T-test Findings for 'him'

	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
him	.183	.669	.737	750	.462	.061	.083

The third person object case singular pronoun's frequency was roughly equivalent across both corpora.

There is no statistically significant difference in use of *him* between L1 (M=0.53, SD=1.12) and L2 (M=0.47, SD=1.16) writing ($t(750.000) = 0.737, p > .05$).

His.

Table 56

Frequencies of 'his'

	L1 English	L2 English
Total Instances	1,034 (0.342%)	1,141 (0.294%)
<u>Adjective Expressions</u>		
possessive adjective	1,034	1,141

Table 57

Group statistics for 'his'

His	Corpus	N	Mean	Std. Deviation	Std. Error Mean
Possessive Adjective	L1 English	380	2.72	4.182	0.215
	L2 English	372	1.59	2.665	0.138

Table 58

T-test Findings for 'his'

his	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
Possessive Adjective Only	33.454	.000	4.418	750	.000	1.132	.256

The third person singular masculine possessive pronoun's frequency was roughly equivalent across both corpora. It occurred only as a possessive adjective, never as a possessive pronoun in either corpus.

There is a statistically significant difference in use of *his* as a possessive adjective L1 (M=2.72, SD=4.18) and L2 (M=1.59, SD=2.66) writing ($t(750.000)=4.418, p<.05$).

Hers.

Table 59

Frequencies of 'hers'

	L1 English	L2 English
Total instances /as % of total corpus	2 (0.001%)	0 (0.000%)
Nominative pronoun	2	0
object	2	0

Table 60

Group Statistics for 'hers'

	Corpus	N	Mean	Std. Deviation	Std. Error Mean
hers	L1 English	380	0.01	0.072	0.004
	L2 English	372	0	0	0

Table 61

T-test Findings for 'hers'

	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
hers	7.936	.005	1.401	750	.162	.005	.004

The third person nominal possessive pronoun is barely present in the L1 English corpus and not at all in the L2 English corpus. There is no statistically significant difference in use of *hers* between L1 (M=0.01, SD=0.07) and L2 (M=0.00, SD=0.00) writing ($t(750.000) = 1.401, p > .05$).

She.

Table 62

Frequencies of 'she'

	L1 English	L2 English
Table 62 (cont'd)		
Total Instances / as % of Corpus	323 (0.107%)	173 (0.045%)
<u>Nominative Expressions</u>	323	173
Argument		
subject	323	173

Table 63

Group Statistics for 'she'

	Corpus	N	Mean	Std. Deviation	Std. Error Mean
She	L1 English	380	0.85	2.678	0.137
	L2 English	372	0.47	1.14	0.059

Table 64

T-test Findings for 'she'

	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
she	26.817	.000	2.538	750	.011	.382	.151

The third person feminine subject case singular pronoun was twice as frequent in the L1 English corpus as the L2 English corpus.

There is a statistically significant difference in use of *she* between L1 (M=0.85, SD=2.67) and L2 (M=0.47, SD=1.14) writing ($t(750.000) = 2.538, p < .05$).

His/her.

Table 65

Concordance Analysis for 'his/her'

	L1 English	L2 English
Total Instances	0	41
as % of total corpus	0.000%	0.011%
<u>Adjective Expressions</u>		
possessive adjective	0	41

The gender-neutral possessive adjective pronoun hedging device occurs 41 times in the L2 English corpus, but never in the L1 English corpus.

There were no surprising differences in usage of the third persons nominative and possessive pronouns between the two corpora. The much greater occurrence in the L1 English corpus of these pronouns, especially the feminine pronouns, is due to the news orientation of *The Economist*. As a newspaper, *The Economist* more often discusses specific political personalities and the much greater occurrence of the feminine pronouns is due to the large role played by Hillary Clinton and Theresa May in the news in 2016 and 2017. The construction *his/her* or *his* or *her* occurs only in the L2 English corpus and reflects the traditional teaching about achieving gender neutrality in English writing. There seems to be momentum at the moment for using the plural form *they* as a gender neutral singular form, although *The Economist Style Guide* does not agree:

If you believe it is “exclusionary” or insulting to women to use he in a general sense, you can rephrase some sentences in the plural. Thus Instruct the reader without lecturing him may be put as Instruct readers without lecturing them. But some sentences resist this treatment: Find a good teacher and take his advice is not easily rendered gender-neutral. So do not be ashamed of sometimes using man to include women, or making he do for she

Not surprisingly, exophoric referencing in was found to be non-existent in either body of writing; cataphoric referencing was found to be infrequent in both corpora (less than one-percent of all references), although slightly more common as a percentage in the L1 English corpus. Cataphoric referencing requires a strong command of style and a knowledge of the readers' expectations, and often involves fronting a dependent clause, which may account for its less frequent occurrence in the L2 English corpus. Cataphoric references were a very small percentage (1% or less) of either corpus and were therefore not reported

CHAPTER 5: CONCLUSIONS

Introduction

This study investigated and described the ways in which L1 English writers and L2 English writers use pronominal referencing to achieve cohesion in their writing. The study described differences and similarities in the relative frequency and proportions of referencing devices in a text. To this end, the study addressed the following research question:

- How do expert L1 English expert writers and Turkish writers of L2 English differ in their use of pronominal and demonstrative reference?

This chapter is composed of four sections. In the first, the findings of the study vis-à-vis each pronoun are evaluated and discussed with respect to the research question. In the second, pedagogical implications are presented. In the third, the limitations of the study are defined. Finally, in the fourth section, areas and directions for further research are suggested.

Findings and Discussion

L2 English writers differ in their use of referent pronouns in some significant ways. Certain pronouns are heavily favored by L1 English writers in their efforts to achieve cohesion across clauses, sentences and texts. They also tend to use certain pronouns in subject and object cases, where L1 English writers do not make such a distinction.

Use of third person impersonal pronoun *it*

Unclear references. Consistent with the findings of the current literature, L2 English writers face challenges in deploying *it* as a cohesive device. L2 English writers display a tendency to use *it* to refer to more than just a singular inanimate object; they seem to be using *it* for extended reference, i.e. to a clause or multiple clauses. Effective extended reference is an essential feature of cohesion in English texts (Halliday & Hassan, 1976). For extended reference L1 English writers where a NS writer might be inclined to use a demonstrative pronoun or a noun phrase. Cho and Shin (2014) reported similar results among Korean EFL students, who, “regardless of their proficiency level, used the pronoun *it* to refer to an extended set of text in the preceding discourse, whereas the pronoun produced by the native group most frequently referred to a nominal item” (p.45).

Some occurrences of unclear reference suggest an attempt to use the pronoun to refer to a clause, several clauses, or even the theme of a paragraph. Similar to Such references are inconsistent with conventional English usage and may derail the cohesion of the writing, leaving the reader confused.

Fronted structures. Perhaps surprisingly, L2 English writers were also much more likely to construct a fronted or cleft structure in their writing (26% of constructions with *it* compared to 17% in the expert corpus). This is a positive finding in that these structures help support back-endedness typical of English writing. These fronted structures however were not always successfully achieved in the L2 English corpus. Constructions such as “It may arise some problems” and “It is underestimated around the world that malnutrition is a big threat to health” and “...it is likely to assert that there is a major utilization of stereotypes and social engineering” were coded as cleft sentences or fronted expressions, despite their strict

syntactic failure. These cases represent an overuse of the fronted structure, where a more direct S-V-O construction would be more concise.

Idiomatic expressions. Idiomatic and non-referential expressions appear frequently in the L2 English corpus. L2 English writers appear to be twice as likely to use idiomatic *it* expressions. However, careful examination of the source files reveals that these idiomatic expressions appear to be clustered in one or two out of nine sections and by essay cycle, suggesting that the students were exposed to the structure in their readings and recycled the structure in their attempts to paraphrase. Still, it is a positive sign that students tried to apply new structures in their own writing.

The third person non-gendered pronoun occurs in the L2 significantly more frequently in the L2 English corpus than in the L1 English corpus and is by far the most problematic of reference pronouns. The findings suggest that L2 overuse the pronoun *it*, although they are much less likely to use the pronoun in the object case (17% of all occurrences compared to 23% for L1 English writers). This difference in itself is not necessarily a threat to cohesion. However, the frequency of unclear or irresolvable references for *it* suggest that Turkish learners of English have difficulty assigning a meaning to the pronoun that the reader can easily decode. Almost all of these problems occurred in subject-position referencing, a finding which differs from what was reported by Cho and Shin (2014, p. 45) among Korean learners who showed “less mastery of using referential expressions in positions other than subject positions”.

L2 English writers’ relative overuse of the pronoun, the tendency to employ it only in a subject position, along with the significant number of unclear references, diminish the overall quality of writing.

Its. The third person impersonal possessive pronoun occurs in the L2 English corpus at one-third the rate as in the L1 English corpus. This underuse suggests L2 English writers may be trying to achieve cohesion through some other (less effective) means: the collocation “of it” occurs 75 times in the L2 English corpus compared to 40 times in the L1 English corpus. This difference seems to reflect L2 English writers’ unfamiliarity with syntactic application of the third person impersonal possessive pronoun.

Demonstrative Reference

This. Turkish L2 English writers use *this* at three times the rate of L1 English writers and represents 0,7% of total words compared to 0,2% of the L1 English corpus. This overuse leads to one of the most problematic areas of textual cohesion, along with the use of *it*, in L2 English writing among native Turkish speakers. This tendency may be due to L2 transfer of Turkish ‘bu’ to represent previously stated ideas. It is worth examining a few examples from the L2 English corpus:

“...because poor countries do not have much choice, they will become unhealthier day by day. They do not usually have the education to know how unhealthy fast food is and this is one of the biggest problems for their health”. In this case, it is unclear whether the writer intends to refer to the lack of education regarding the healthfulness of fast food, or the fast food itself which is the problem. This lack of precision is a persistent problem. In other cases, the problem is not a lack of precision, but a general lack of clarity of reference: “Germany didn't face the reality of citizens which labeled as guest workers till 2000. This is used for social engineering. With this way, German Turks didn't accepted [sic] as equal citizens”.

There is a tendency among native Turkish-speaking L2 English writers to link ideas using the phrase “and this”, where *this* is a pronominal, not adjectival. Such a technique is not frequent in the L1 English corpus (seven occurrences) but is quite frequent in the L2 English corpus (116 occurrences) and seems to represent language transfer. As a technique for textual cohesion in the L2 English corpus, it is rarely effective.

Although the proportion of the adjective for of the pronoun is much larger in the L2 English corpus, its use is less problematic than it is as in its pronominal form.

That. Use of the referring expression *that* was not as problematic as it was for some of the other expressions. It is worth noting however, that Turkish L2 English writers are much less likely to use the term in a subject position and much more likely to use it in an object position. This may reflect the tendency among Turkish speakers to use the demonstrative of close proximity (*this/bu*) in subject positions and the demonstrative adjective of far position (*that/şu*) in object cases. Although Turkish learners might benefit from learning that the term has greater flexibility in English than they realize, cohesion was generally not impeded.

These. The higher incidence of *these* in L2 writing was remarkable. While 24% of occurrences in the L1 English corpus represented a nominal pronoun, only 9% of occurrences in the L2 were nominal pronouns. It is hard to say what accounts for L2 English writers’ preference for using *these* as an adjectival pronoun, but at a rate of seven times the L1 English writers, it is a distinct difference.

Those. All in all, L1 and L2 English writers used this pronoun in roughly equal proportions as a subject, object and complement, and as non-qualified, qualified personal and qualified impersonal pronouns. The remarkable difference was in the proportion of nominal pronouns to adjective pronouns: In the L1 English

corpus, there was an 80/20 preference for nominal pronouns over adjectives. In the L2 English corpus, this proportion was reversed. This, and other findings are similar to those found by Petch-Tyson (2000) in her comparative study of demonstrative referencing among American L1 English writers and L2 English writers from Finland, Sweden and the Netherlands. It is important to note that L1 English writers frequently deploy *those* in a qualified cataphoric referencing structure to reference a particular class of persons, a conventional feature of L1 writing markedly absent from many L2 English corpora.

All of these results for demonstrative referencing (*this that, these and those*) are also consistent with Hinkel's (2001) findings in his study of Japanese, Korean, Indonesian and Arabic speakers, who also used such pronouns to refer to broader contexts or to ideas without clear textual antecedents Likewise, Zarepour (2016) reported that demonstrative reference was a "major weakness" among even advanced Iranian L2 English writers (p. 413).

Implications for Practice

In accordance with the above-described findings, this study offers some implications for the teaching of writing for intermediate students in an academic environment. The importance of cohesion in academic writing cannot be overestimated and yet the focus of cohesion in academic has been by and large restricted to transitions and coordinating conjunctions (Hinkel, 2001, p.112) Instruction of grammatical cohesion, especially through referencing remains appears to be insufficient for Turkish students as has been documented for speakers of many other languages (Hinkel, 2001; Kim, 2012; Leńko-Szymańska, 2004; Petch-Tyson, 2007; Swierzbın, 2010; Yang, 2012; Zhang, 2015).

In some of the findings, it is the relative proportions that tell the story. The sizeable difference (overuse of *it*, *this*, *that*, and *these* indicate that student may be asking these pronouns to do jobs that they are in capable of. Petch-Tyson (2000) refers to these extended or non-nominal references as “situation” reference while Halliday and Hasan (1976) call them “text” reference. The natural transfer of L1 rules can make the achievement of grammatical cohesion elusive.

Because this English L2 population employs pronouns in ways that sometimes fail to effectively unite ideas across their text, explicit instruction in the effective use of pronominal referencing must begin when students begin to write paragraphs. Teachers must be explicit in their instruction as to the kinds of referents that can be represented by referential pronouns: *it* is not insufficient to carry extended referential meaning, and demonstrative pronouns alone are not always sufficient. At the low-to-high intermediate levels, students should focus on learning a range of enumerative nouns that categorize concepts or actions. Such instruction could easily be based on student-produced writing where cohesion was not achieved.

By sharing the piece of writing and asking the students what was represented specifically by the referring pronoun, the teacher can elicit a categorical noun. For example, one student has written: “Humanity has continually decreased violence and this has become exponentially faster in recent eras...”. In this case, *this* is being expected to carry the meaning of a clause. The teacher might ask “what exactly has become ‘exponentially faster’” in order to elicit a noun that conveys the correct meaning. ‘Decrease’ is an obvious answer, but also the generic ‘change’ or the more specific ‘reduction.’ Although somewhat beyond the scope of this study, a quick query of collocations with *this* confirms the author’s experience that shows Turkish learners tend to rely on a very limited range of such generic enumerative nouns, especially ‘situation’ and ‘example’

Although *this* and *that* connote nearness and farness in spoken English, in formal written English, the connotations are somewhat less obvious. Students would be helped to treat *this* as a term that suggests the writer has some involvement or interest, and that *that* can represent ideas from which to maintain some distance or limit the discourse.

At a more advanced level, students might learn words that connote a positive or negative meaning, such as ‘improvement’ or learn to construct a noun phrase to even more precisely convey the writer’s point of view, for example ‘this welcome change in human relations’ not only achieves a higher level of cohesion, but also communicates to the reader the writer’s perspective.

The relative frequency of the coordinating expression “and this’ in the L2 English corpus suggests learners are not using subordination as effectively as they might in order to highlight one clause over another in the same sentence. Students might be reminded that construction of an adjective clause with ‘which’ is often a

better solution than a coordinating conjunction and a referencing pronoun that may not be clear.

Student use of *it* in fronted structures is also problematic. 'It is seen that' remains a persistent, awkward and transfer from L1 to L2. Other awkward examples such as 'It mustn't be underestimated that this test won't show a person is a psychopath' and 'it may not be ignored the fact that there are also some young people who cannot get a chance' are a reminder that teachers must focus on simplicity of construction and a preference for an S-V-O orientation, even when the verb is expressed in a passive voice.

The infrequency of the possessive pronoun *its*, along with the over use of the potentially awkward use of *of it* suggests that students should be reminded of the function of this pronoun.

Students of English for academic purposes would also benefit from explicit instruction in the effective use of the nominative pronoun *those* as a catch all for people, especially when combined with a qualifying adjectival expression. That 80% of expressions of *those* were nominal in the L1 English corpus compared to 23% in the L2 English corpus suggests that students are underutilizing this structure. Further indirect evidence comes from corpus searches for the word 'people', (the 10th most frequent word in the L2 English corpus) which occurs 5,080 times in the student corpus but 607 times in the expert corpus.

In general, teachers should make use of models of good writing to raise students' awareness of effective referencing as well as samples of student writing to give students guided practice. Crosswaithe (2017) has devised a set of awareness-raising exercises for EAP students at the University of Hong Kong. These exercises ask student to describe the referring function and scope of pronouns *it* and *this* in an

academic passage. Turkish students would likewise benefit from such exposure and explicit instruction. Student writing may improve when they are shown how expert writers achieve cohesion across a text and are encouraged to emulate those techniques and patterns.

Implications for Further Research

The potential for exploration of the nature of cohesion in L2 academic writing is vast. A deeper investigation into the nature of referential confusion and a more precise understanding of what breaks a referential link could provide valuable pedagogical insight. Assuming L1 influences a writer's L2 production, a comparative analysis of pronominal and demonstrative referencing would shed light on areas of difference. Knowing how L2 English writers make decisions about referring expressions will allow educators to guide students toward more conventional or standard uses of referencing for cohesion that satisfy readers expectations.

Some standardization of methods is also called for. The nature of both lexical and grammatical cohesion is complex, and researchers have taken many approaches to study the subject. The best approaches should be replicated across the L1 range of first languages to determine to what extent current findings are generalizable. Languages. This study had intended to examine and code for the type of the or head reference for each of the pronouns (singular or plural noun, clause, multiple clauses). Future research might categorize and standardize for analysis the types of 'extended' referents used in academic English. This analysis of pronominal referencing of Turkish L2 academic writing may serve as baseline study against which repeat studies among speakers of other languages, especially Chinese, Korean, Japanese and Farsi may offer insights into the differences among learners.

It may also be useful to investigate why certain generic nouns occur so much more frequently in L2 writing, specifically 'people'. Such a study might involve a corpus of L2 writing of speakers of other languages to discover whether Turkish L2 English writers exhibit a demonstrable preference for personalized over conceptual references (e.g., mentally-ill people over mental illness; Russians over Russia) than

other L2 English writers. An experimental study in which the pedagogical implications mentioned above are applied in a classroom setting with a control group would contribute to the development of actionable strategies to improve cohesion in L2 academic writing.

Limitations of the Study

There are several limitations to this study that suggest that the findings must be interpreted with caution. The corpus of student writing was the product of convenience sampling and is composed of essays that were written on a limited range of topics and within the same institution. If the corpus were expanded to include students from other institutions, writing about different topics, the results may have been different. Second, the expert corpus is likewise the result of convenience sampling, and as a body of writing concerned with political issues, may exhibit some differences which could limit the value of the comparison with the learner corpus.

It is possible that the differences observed between the use of referencing between native and non-native writers may be the results of professional and non-professional styles. It would be worthwhile to complement this study by a comparing learners' compositions with those written by non-professional native writers (first-year university students) in order to know whether overuse of the *it*, *this*, *that*, *these*, *those*, *they*, and *them* is associated with non-nativeness or a lack of expertise.

The L2 English writers in this study also wrote their essays in class, albeit over several days, which may introduce environmental differences in the setting for production. Differences in motivation and age might result in differences in performance, limiting the compatibility of the corpora. The study also relied on the judgment of one rater to assign codes for parts of speech, argument and other features. There is undoubtedly a chance for human error. Because of time constraints, it was not possible to tag the pronouns in the source files for their function, rather they were coded in a spreadsheet generated from a corpus query tool.

Conclusion

This study was conceived as a means to better understand how Turkish L2 English learners use referencing to achieve cohesion in their essays compared to expert writers. It presents a quantitative and qualitative assessment of the use of referential pronouns in NS and NNS writing and their contribution to textual cohesion. The research has demonstrated that L2 English writers use referential pronouns differently from expert writers. In particular, L2 English writers employ the pronouns *it*, *this*, *that*, *these*, *those*, and *they* to a far greater extent than, and qualitatively differently from, L1 expert writers. Of the 32 salient referential devices (which excludes the personal gender-specific pronouns and which excludes *this* and *that* in non-referential structures, i.e., noun clauses, adverbials and adjective clauses) examined in this study, 27 had statistically significant occurrences between the two corpora. In other words, Turkish writers of L2 English consistently use pronominal referencing in significantly different ways from L1 English expert writers.

This overuse and, less frequently, underuse, is likely a result of L1 transfer, whereby referencing pronouns are expected to carry meaning that is not consistent with the norms of written English. At best, the result is unnatural-sounding; at worst, coherence is lost. These patterns are similar to those observed in previous studies showing that L2 English learners overuse discourse markers to achieve cohesion (Aysu, 2017; Öztürk & Köse, 2016).

The construction of a reader-centric essay with a unified flow of information is a major challenge for L2 English writers, given that argumentation requires the writers to establish clear and unambiguous logical links of abstract ideas across clauses, sentences and paragraphs. Proper use of English demonstratives to achieve cohesion is a persistent a problem due not only to their highly idiosyncratic nature of

their applications but also to a lack of explicit instruction. Still, there are opportunities to improve L2 English writers' competence. By abandoning a reductionist approach to teaching superficial parallel meaning of English demonstrative pronouns to beginners, by stressing the importance of proximity in pronoun resolution to low-intermediate students, and by focusing on the range of anaphora associated with referent classes to high-intermediate learners, educators can empower L2 English writers to successfully communicate their ideas and to raise the quality of their writing.

REFERENCES

- Abbott, D. (2017). Are you looking for an argument? *Modern English Teacher*, (1), 34. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=edsgao&AN=edsgcl.509451282&site=eds-live>
- Alarcon, J. B., & Morales, K. N. (2011, June). Grammatical cohesion in students' argumentative essay. *Journal of English and Literature*, 2(5), 114-127. Retrieved from [http://www.academicjournals.org/article/article1379417099_Alarcon and Morales.pdf](http://www.academicjournals.org/article/article1379417099_Alarcon%20and%20Morales.pdf)
- Aldera, A. (2016). Cohesion in written discourse: A case study of Arab EFL students. *Arab World English Journal*, 7(2,), 328–341. Retrieved from doi:10.24093/awej/vol7no2.22.
- Aktas, R. N., & Cortes, V. (2008). Shell nouns as cohesive devices in published and ESL student writing. *Journal of English for Academic Purposes*, 7, 3-14. 10.1016/j.jeap.2008.02.002 Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=edselp&AN=S1475158508000039&site=eds-live>
- Alshahrani, A. A. S. (2015). A cross-linguistic analysis of interactive metadiscourse devices employment in native English and Arab ESL academic writings. *Theory and Practice in Language Studies*, 5(8), 1535. 10.17507/tppls.0508.01 Retrieved

from <https://search.proquest.com/docview/1706203523> <https://search.proquest.com/docview/1706203523>

Anderson, G. (2014). So, transitions: Linking adverbial use of university ESL students. *CATESOL Journal*, 26(1), 1-13. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=eue&AN=110326580&site=eds-live>

Aysu, S. (2017). The use of discourse markers in the writings of Turkish students of English as a foreign language: A corpus based study. *Journal of Higher Education & Science / Yükseköğretim Ve Bilim Dergisi*, 7(1), 132. doi:10.5961/jhes.2017.191

Bacha, N. N. (2010). Teaching the academic argument in a university EFL environment. *Journal of English for Academic Purposes*, 9(3), 229-241. 10.1016/j.jeap.2010.05.001 Retrieved from <http://www.sciencedirect.com/science/article/pii/S1475158510000378>

Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: investigating language structure and use*. New York: Cambridge University Press, 1998. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=cat00040a&AN=bilk.306911&site=eds-live>

Biber, D., Conrad, S., & Reppen, R. (2001). *Corpus linguistics and English for specific purposes* (Digital print ed.). Cambridge: Cambridge University Press.

Breeze, R. (2008). Researching simplicity and sophistication in student writing. *International Journal of English Studies (IJES)*, 8(1) 51-66.

- Bonelli, E. T. (2010). Theoretical overview of the evolution of corpus linguistics. In A. O'Keeffe, & M. McCarthy (Eds.), *The Routledge Handbook of Corpus Linguistics* (pp. 14-28). London: Routledge.
- Brown, G., & Yule, G. (1996). *Discourse analysis*. Cambridge: Cambridge University Press.
- Bychkovska, T., & Lee, J. J. (2017). At the same time: Lexical bundles in L1 and L2 university student argumentative writing. *Journal of English for Academic Purposes*, 30, 38-52. 10.1016/j.jeap.2017.10.008 Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=edselp&AN=S1475158517301005&site=eds-live>
- Carroll, J., & Dunkelblau, H. (2011). Preparing ESL students for "real" college writing: A glimpse of common writing tasks ESL students encounter at one community college. *Teaching English in the Two Year College*, 38(3), 271. Retrieved from http://gateway.proquest.com/openurl?ctx_ver=Z39.88-2003&xri:pqil:res_ver=0.2&res_id=xri:ilcs-us&rft_id=xri:ilcs:rec:abell:R04397515
- Cennetkusu, N. G. (2017). International students' challenges in academic writing: A case study from a prominent U.S. university. *Journal of Language & Linguistics Studies*, 13(2), 309-323. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=eue&AN=126127409&site=eds-live>
- Chandrasegaran, A. (2008). NNS students' arguments in English: Observations in formal and informal contexts. *Journal of Second Language Writing*, 17(4), 237-254. doi:10.1016/j.jslw.2008.04.003

- Coksun, E. (2011). Cohesion in compositions of Turkish and immigrant students. *Educational Sciences: Theory & Practice*, 11(2), 892-899. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=eue&AN=61029630&site=eds-live>
- Connor, U. (1990). Linguistic/rhetorical measures for international persuasive student writing. *Research in the Teaching of English*, 24(1), 67-87. Retrieved from <http://www.jstor.org/stable/40171446> <http://www.jstor.org/stable/40171446>
- Conrad, S. (2010). What can a corpus tell us about the grammar? In A. O’Keeffe, & M. McCarthy (Eds.), *The routledge handbook of corpus linguistics* (pp. 227-240). London: Routledge.
- Crossley, S. A., Kyle, K., & McNamara, D. S. (2016). The development and use of cohesive devices in L2 writing and their relations to judgments of essay quality. *Journal of Second Language Writing*, 32, 1-16.
doi:10.1016/j.jslw.2016.01.003
- Crossley, S. A., & McNamara, D. S. (2009). Computational assessment of lexical differences in L1 and L2 writing. *Journal of Second Language Writing*, 18(2), 119-135. doi:10.1016/j.jslw.2009.02.002
- Crossthwaite, P. d. (2016). Managing referential movement in Asian L2 writing: Implications for pedagogy. *Writing & Pedagogy*, 8(3), 537-558.
doi:10.1558/wap.27695
- The Economist style guide*. (2010). London: Profile Books.
- Flowerdew, L. (2012). *Corpora and language education*. New York, NY : Palgrave Macmillan, 2012. Retrieved

from <http://search.ebscohost.com/login.aspx?direct=true&db=cat00040a&AN=bilk.830993&site=eds-live>

Flowerdew, L. (2015). Corpus-based research and pedagogy in EAP: From lexis to genre. *Language Teaching*, 48(1), 99. 10.1017/S0261444813000037 Retrieved from <https://search.proquest.com/docview/1643372598> <https://search.proquest.com/docview/1643372598>

Gao, X. (2016). A cross-disciplinary corpus-based study on English and Chinese native speakers' use of linking adverbials in academic writing. *Journal of English for Academic Purposes*, 24, 14-28. 10.1016/j.jeap.2016.08.002 Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=edselp&AN=S147515851630056X&site=eds-live>

Ghasemi, M. (2013). An investigation into the use of cohesive devices in second language writings. *Theory and Practice in Language Studies*, 3(9), 1615. 10.4304/tpls.3.9.1615-1623 Retrieved from <https://search.proquest.com/docview/1445143145> <https://search.proquest.com/docview/1445143145>

Gee, J. P. (2014). *The Routledge handbook of discourse analysis*. London: Routledge.

Gray, B. (2010). On the use of demonstrative pronouns and determiners as cohesive devices: A focus on sentence-initial this/these in academic prose. *Journal of English for Academic Purposes*, 9(3), 167-183. 10.1016/j.jeap.2009.11.003 Retrieved from <http://www.sciencedirect.com/science/article/pii/S1475158509000678>

- Gray, B. (2015). *Linguistic variation in research articles*. Amsterdam ; Philadelphia: John Benjamins Publishing Company. Retrieved from <http://www.sciencedirect.com/science/article/pii/S1475158509000678> <http://www.sciencedirect.com/science/article/pii/S1475158509000678>
- Green, C. (2012). A computational investigation of cohesion and lexical network density in L2 writing. *English Language Teaching*, 5(8), 57-69. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ1079753&site=eds-live>
- Güngör, F., & Uysal, H. H. (2016). A comparative analysis of lexical bundles used by native and non-native scholars. *English Language Teaching*, 9(6), 176-188. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ1101230&site=eds-live>
- Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. London : Longman, 1976 (imp. 98); 1st ed., 9th imp. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=cat00040a&AN=bilk.856281&site=eds-live>
- Hinkel, E. (2001). Matters of cohesion in L2 academic texts. *Applied Language Learning*, 12(2), 111-32. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ648973&site=eds-live>
- Hinkel, E. (2005a). *Handbook of research in second language teaching and learning*. Mahwah, N.J.: Routledge. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=128558&site=eds-live>

- Hinkel, E. (2005b). Hedging, inflating, and persuading in L2 academic writing. *Applied Language Learning*, 15(1-2), 29. Retrieved from http://gateway.proquest.com/openurl?ctx_ver=Z39.88-2003&xri:pqil:res_ver=0.2&res_id=xri:ilcs-us&rft_id=xri:ilcs:rec:abell:R03829641
- Hinkel, E. (2009). *Teaching academic ESL writing practical techniques in vocabulary and grammar*. London: Routledge.
- Kafes, H. (2012). Lexical cohesion: An issue only in the foreign language? *English Language Teaching*, 5(3), 83-94. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ1078971&site=eds-live>
- Kehoe, A., & Renouf, A. (2009a). *Corpus linguistics : Refinements and reassessments*. Amsterdam: Brill Academic Publishers. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=281504&site=eds-live>
- Kim, Y. (2012). An analysis of referential use in Korean EFL learners' argumentative essays. *English Teaching*, 67(3), 181-204. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=mzh&AN=2014650789&site=eds-live>
- Leńko-Szymańska, Agnieszka, S. (2009). Demonstratives as anaphora markers in advanced learners' English. In G. Aston, S. Bernadini, & S. Stewart (Eds.), *Coprorra and Language Learners* (pp. 89-107) Philadelphia, PA: John Benjamins.

- Leo, K. (2012). Investigating cohesion and coherence discourse strategies of Chinese students with varied lengths of residence in Canada. *TESL Canada Journal*, 29, 157.
- Liu, M., & Braine, G. (2005). Cohesive features in argumentative writing produced by Chinese undergraduates. *System*, 33(4), 623-636.
 //doi.org/10.1016/j.system.2005.02.002 Retrieved
 from <http://www.sciencedirect.com/science/article/pii/S0346251X05000692>
- McCarthy, D., & Sampson, G. (2005). *Corpus linguistics: Readings in a widening discipline*. London: Continuum. Retrieved
 from <http://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=377886&site=eds-live>
- McCarthy, M., & O'Keeffe, A. (2010). *The routledge handbook of corpus linguistics* Routledge.10.4324/9780203856949 Retrieved
 from <http://dx.doi.org/10.4324/9780203856949>
- McEney, T., & Hardie, A. (2012). *Corpsu Linguistics: Method, theory and practice*. New York, NY: Cambridge University Press.
- McEney, T., Xiao, R., & Tono, Y. (2006). *Corpus-based language studiesf: An advanced resource book*. London: Routledge.
- McNamara, D. S., Crossley, S. A., & McCarthy, P. M. (2010). Linguistic features of writing quality. *Written Communication*, 27(1), 57-86. Retrieved
 from <http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ876566&site=eds-live> <http://dx.doi.org/10.1177/0741088309351547>
- Meisuo, Z. (2000). Cohesive features in the expository writing of undergraduates in two Chinese universities. *RELC Journal*, 31(1), 61-95.
 10.1177/003368820003100104 Retrieved

from <http://search.ebscohost.com/login.aspx?direct=true&db=edselc&AN=edselc.2-52.0-27744465727&site=eds-live>

Meyer, C. F. (2002). *English Corpus Linguistics : An Introduction*. Cambridge, UK: Cambridge University Press. Retrieved

from <http://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=112619&site=eds-live>

Moon, R. (2010). What can a corpus tell us about lexis? In A. O'Keeffe, & M. McCarthy (Eds.), *the routledge handbook of corpus linguistics* (pp. 197-211). London: Routledge.

Naderi, S., Keong, Y. C., & Latif, H. (2013). The use of referential cohesion in academic texts by Persian EFL learners. *GEMA Online Journal of Language Studies*, 13(3), 45-62. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=edselc&AN=edselc.2-52.0-84884785386&site=eds-live>

Neff, J., Ballesteros, F., Dafouz, E., Martinez, F., Rica, J., Diez, M., & Prieto, R. (2007). A contrastive functional analysis of errors in Spanish EFL university writers' argumentative texts: Corpus-based study. *Language & Computers*, 60(1), 203-225. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=26863903&site=ehost-live>

Nunan, D. (1993). *Introducing discourse analysis*. London. Penguin Books.

O'Halloran, K., Hewings, A., Coffin, C., & University, O. (2004). *Applying english grammar: Corpus and functional approaches*. London: Routledge. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=790351&site=eds-live>

- Öztürk, Y., & Köse, G. D. (2016). Turkish and native English academic writers' use of lexical bundles. *Journal of Language and Linguistic Studies*, 12(1), 149-165.
- Palmer, J. C. (1999). Coherence and cohesion in the English language classroom: The use of lexical reiteration and pronominalisation. *RELC Journal*, 30(2), 61. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=edb&AN=53581409&site=eds-live>
- Park, E. (2016). Issues of international students' academic adaptation in the ESL writing class: A mixed-methods study. *Journal of International Students*, 6(4), 887. Retrieved from <https://search.proquest.com/docview/1861791241> <https://search.proquest.com/docview/1861791241>
- Park, S. (2013). Lexical analysis of Korean university students' narrative and argumentative essays. *English Teaching*, 68(3), 7. doi:10.15858/engtea.68.3.201309.131.
- Park, Y. (2013). How Korean EFL students use conjunctive adverbials in argumentative writing. *English Teaching*, 68(4), 263-284. doi:10.15858/engtea.68.4.201312.263
- Petch-Tyson, S. (2000). Demonstrative expressions in argumentative discourse: A computer corpus-based comparison of non-native and native English. In S. Botley & AM Anthony (Eds.), *Corpus-based and computational approaches to discourse anaphora demonstrative expressions* (pp. 43-64). Philadelphia, PA: John Benjamins.

- Petchprasert, A. (2013). A study of cohesive markers used in L1 and L2 essay writing: Translation versus direct composition. *3L: Southeast Asian Journal of English Language Studies*, 19(1), 19-33. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=eue&AN=94996707&site=eds-live>
- Renouf, A., & Kehoe, A. (2006). *The changing face of corpus linguistics*. Amsterdam: Brill Academic Publishers. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=160191&site=eds-live>
- Roemer, U., & Roemer, U. (2011). *Corpus research applications in second language teaching*. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=edswah&AN=000295464300012&site=eds-live>
- Rowley-Jolivet, E., Carter-Thomas, S., & Boulton, A. (2012). *Corpus-informed Research and Learning in ESP: Issues and Applications*. Amsterdam: John Benjamins Publishing Company. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=451547&site=eds-live>
- Ryan, J. (2015). Overexplicit referent tracking in L2 English: Strategy, avoidance, or myth? *Language Learning*, 65(4), 824-859. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=mzh&AN=2016930017&site=eds-live>
- Sabet, M. K., & Minaei, R. (2017). A comparative corpus-based analysis of genre specific discourse: The quantitative and qualitative academic papers in the field of the TEFL. *Theory and Practice in Language Studies*, 7(4), 294.

10.17507/tpis.0704.08 Retrieved

from <https://search.proquest.com/docview/1888920188>
<https://search.proquest.com/docview/1888920188>

Schneer, D. (2014). Rethinking the argumentative essay. *TESOL Journal*, 5(4), 619-653. 10.1002/tesj.123 Retrieved

from <http://onlinelibrary.wiley.com/doi/10.1002/tesj.123/abstract>
<http://onlinelibrary.wiley.com/doi/10.1002/tesj.123/abstract>

Stab, C., & Gurevych, I. (2014). Identifying argumentative discourse structures in persuasive essays. Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP). doi:10.3115/v1/d14-1006

Swierzbina, B. (2010). Demonstratives' special place in the English reference system:

Why that's important for English language learners. *Language & Linguistics Compass*, 4(10), 987. Retrieved from

<http://search.ebscohost.com/login.aspx?direct=true&db=edb&AN=66591071&site=eds-live>

Thornbury, S. (2010). What can a corpus tell us about discourse? In A. O'Keeffe, & M. McCarthy (Eds.), *The routledge handbook of corpus linguistics* (pp. 270-287). London: Routledge.

Tognini-Bonelli, E. (2001). *Corpus linguistics at work*. Amsterdam: John Benjamins Publishing Co. Retrieved

from <http://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=363353&site=eds-live>

Uçar, S., & Yükselir, C. (2017). A corpus-based study on the use of the logical connector "thus" in the academic writing of Turkish EFL learners. *English Language Teaching*, 10(2), 64-72. Retrieved

from <http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ1126745&site=eds-live>

Uysal, H. H. (2008). Tracing the culture behind writing: Rhetorical patterns and bidirectional transfer in L1 and L2 essays of Turkish writers in relation to educational context. *Journal of Second Language Writing, 17*, 183-207.

10.1016/j.jslw.2007.11.003. Retrieved

from <http://search.ebscohost.com/login.aspx?direct=true&db=edselp&AN=S106037430700080X&site=eds-live>

Wang, Y. (2012). Differences in L1 and L2 academic writing. *Theory and Practice in Language Studies, 2*(3), 637. Retrieved

from http://gateway.proquest.com/openurl?ctx_ver=Z39.88-

[2003&xri:pqil:res_ver=0.2&res_id=xri:ilcsus&rft_id=xri:ilcs:rec:abell:R04818356](http://gateway.proquest.com/openurl?ctx_ver=Z39.88-2003&xri:pqil:res_ver=0.2&res_id=xri:ilcsus&rft_id=xri:ilcs:rec:abell:R04818356)

Wind, S. A., Stager, C., & Patil, Y. J. (2017). Exploring the relationship between textual characteristics and rating quality in rater-mediated writing assessments: An illustration with L1 and L2 writing assessments. *Assessing Writing, 34*, 1-15. 10.1016/j.asw.2017.08.003 Retrieved

from <http://search.ebscohost.com/login.aspx?direct=true&db=edselp&AN=S1075293517300247&site=eds-live>

Witte, S. P., & Faigley, L. (1981). Coherence, cohesion, and writing quality. *College Composition and Communication, 32*(2), 189-204. 10.2307/356693 Retrieved

from <http://www.jstor.org/stable/356693>

Yagız, O. (2016). The academic writing approaches of Turkish graduate students of English language. *Journal of Graduate School of Social Sciences, 20*(1), 73-92.

Retrieved

from <http://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=115175610&site=eds-live>

- Yang, W., & Sun, Y. (2012). The use of cohesive devices in argumentative writing by Chinese EFL learners at different proficiency levels. *Linguistics and Education*, 23(1), 31-48. 10.1016/j.linged.2011.09.004. Retrieved from <http://www.sciencedirect.com/science/article/pii/S0898589811000842>
- Yilmaz, E., & Dikilitas, K. (2017). EFL learners' uses of adverbs in argumentative essays. *Novitas-ROYAL*, 11(1), 69-87. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=eue&AN=123533340&site=eds-live>
- Yang, W., & Sun, Y. (2012). The use of cohesive devices in argumentative writing by Chinese EFL learners at different proficiency levels. *Linguistics and Education*, 2331-48. doi:10.1016/j.linged.2011.09.004
- Yilmaz, E., & Dikilitas, K. (2017). EFL learners' uses of adverbs in argumentative essays. *Novitas-ROYAL*, 11(1), 69-87.
- Zarepour, F. (2016). Cohesion analysis of Iranian advanced EFL learners' writing. *Journal of Language Teaching and Research*, (2), 408. 10.17507/jltr.0702.23 Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=edsgao&AN=eds-gcl.461704511&site=eds-live>
- Zhang, J. (2015). An analysis of the use of demonstratives in argumentative discourse by Chinese EFL learners. *Journal of Language Teaching and Research*, 6(2), 460. 10.17507/jltr.0602.29