REPUCLIC OF TURKEY
ÇUKUROVA UNIVERSITY
INSTITUTE OF SOCIAL SCIENCES
DEPARTMENT OF ENGLISH LANGUAGE TEACHING

THE USE OF RELATIVE CLAUSES WITH THE RELATIVE PRONOUN *WHICH* IN TURKISH UNIVERSITY STUDENTS' WRITTEN ENGLISH

Mustafa ÇOBAN

MASTER OF ARTS

ADANA / 2011

REPUCLIC OF TURKEY
ÇUKUROVA UNIVERSITY
INSTITUTE OF SOCIAL SCIENCES
DEPARTMENT OF ENGLISH LANGUAGE TEACHING

THE USE OF RELATIVE CLAUSES WITH THE RELATIVE PRONOUN *WHICH* IN TURKISH UNIVERSITY STUDENTS' WRITTEN ENGLISH

Mustafa ÇOBAN

Supervisor: Assist. Prof. Dr. Cem CAN

MASTER OF ARTS

ADANA /2011

ÖZET

HANGİ ORTAÇ ADILI İLE ORTAÇ TÜMCECİKLERİNİN TÜRK ÜNİVERSİTE ÖĞRENCİLERİN YAZILI İNGİLİZCESİNDE KULLANIMI

Mustafa ÇOBAN

Yüksek Lisans Tezi, İngiliz Dili Eğitimi Anabilim Dalı
Kasım 2011, 69 sayfa

Öğrenici Derlemi, öğrenicilerin aradil gelişimini farklı dil grupları arasında karşılaştırma olanağı sağlaması bakımından gelişen bir araştırma alanıdır. Can (2010, s.16) Uluslararası Öğrenici Derlemi'nin (ICLE), bu alanlara aradil gelişimini etkileyen çeşitli değişkenleri incelemeye yarayacak bol miktarda veri sağladığını belirtmektedir. Anadil ve aradil karşılaştırmaları öğrenici dilinin anadile benzemeyen yönlerini özellikle ileri düzey öğrencilerde belirli kelime veya dilbilgisel yapıların kullanım sıklığını ortaya çıkarmayı amaçlar (Granger, 1998, s.12).

Bağımlı tümcecikler dilbilimsel açıdan karmaşık yapılardan birisi olarak değerlendirilmektedir. Bağımlı tümcecik yapılarının içerisinde ortaç tümcecikleri yazılı metinlerde daha çok bulunmaktadır. Biber (2006, s. 76) bağımlı tümcecikler içerisindeki ortaç tümceciklerinin öğrenicilerin özellikle yazılı metinlerde karşılaştığı başlıca zorlayıcı dilbilimsel yapılardan birisi olduğunu ileri sürmektedir. Baş konumundaki ad öbeğinin gönderimini belirlemek ve daha detaylı bilgi sağlamak ortaç tümceciklerin başlıca iki temel işlevi olarak adlandırılır. Bununla birlikte, birçok durum içinde *Which* ortaç adılı ile ortaç tümcecikleri bu iki temel işlevi karşılamaktadır.

Bu çalışmanın amacı *Which* ortaç adılı ile ortaç tümceciklerinin Türk üniversite öğrenicilerin yazılı İngilizcesindeki kullanımını Amerikalı üniversite öğrencilerinin ve yabancı dil olarak yazılı İngilizce derlemlerinde incelemektir. Bu çalışma ile Türk üniversite öğrenicilerin *Which* ortaç adılı ile ortaç tümceciklerini kullanımlarının anadili İngilizce olanlarla ne ölçüde benzerlik gösterdiği; *Which* ortaç adılı ile ortaç tümceciklerin kullanım sıklığını ve *Which* ortaç adılı ile ortaç tümceciklerin kullanımında olası bir anadil etkisini araştırmaktır. Uluslararası Öğrenici İngilizcesi Derlemi (ICLE) ve alt derlemleri Uluslararası Türk Öğrenici İngilizcesi Derlemi (TICLE), Uluslararası Japon Öğrenici İngilizcesi Derlemi (JPICLE), Louvain Anadili

İngilizce Olanların Komposizyonları Derlemi (LOCNESS) ve Türk Üniversite Derlemi (TUC) başlıca çalışma materyalleri olarak kullanılmıştır.

Anahtar Kelimeler: Ortaç Tümcecikler, Derlem Bilim, *Which* ortaç adılı ile ortaç tümcecikleri, Aradil, Karşıtsal Aradil Çözümlemesi, Öğrenici Derlemi

ABSTRACT

THE USE OF RELATIVE CLAUSES WITH THE RELATIVE PRONOUN *WHICH* IN TURKISH UNIVERSITY STUDENTS' WRITTEN ENGLISH

Mustafa ÇOBAN

MA Thesis, English Language Teaching Department
Supervisor: Asst. Prof. Dr. Cem CAN
November, 2011, 69 pages

Learner Corpus has become a growing research field providing the chance of comparing and contrasting native and non-native speakers of a language in comparable materials. Can (2010, p.16) states that a learner corpus, like International Corpus of Learner English (ICLE), contributes to language studies with the abundance of hands-on material enabling the researchers to examine various variables that could affect English interlanguage. Granger (1998, p.12) suggests that native language and interlanguage comparisons aim to uncover the features of non-nativeness of learner language.

Dependent clauses are often considered to be a type of linguistic complexity and as a result they have been associated with writing rather than speech. In the types of dependent clauses, relative clauses are much more common in the written texts. According to Biber (2006, p. 76) relative clauses have two primary functions: to specify the reference of the head noun, or to provide elaborating information. However, in many situations the relative clauses with the relative pronoun *which* serves both functions .

This particular study aims at examining the use of relative clauses with the relative pronoun *which* in Turkish EFL learners' argumentative essays with the use in argumentative essays of native speaker American university students, Japanese EFL learners' and native Turkish argumentative essays. The study addresses the questions of whether Turkish learners use the relative pronoun *which* to the same extent as native speaker American university students as regards to the frequency of overuse and

underuse and seeks the possible mother tongue effect. In this study, Turkish International Corpus of Learner English (TICLE), Japanese International Corpus of Learner English (JPICLE), which are the sub corpora of International Corpus of Learner English (ICLE), Louvain Corpus of Native English Essays (LOCNESS) and Turkish University Corpus (TUC) are used as principal materials.

Keywords: Relative Clauses, Corpus Linguistics, Contrastive Interlanguage Analysis, Interlanguage, Learner Corpus.

ACKNOWLEDGMENTS

First and foremost I want to thank my advisor Asst. Prof. Dr. Cem CAN. It has been an honor to be his master student. He has taught me, both consciously and unconsciously, how well a corpus linguistics analysis is done. I appreciate all his contributions of time, ideas, and funding to make my MA degree experience productive and stimulating. The joy and enthusiasm he has for his language studies was contagious and motivational for me. I am also thankful for the excellent example he has provided as a successful academician.

I would like to express my sincere thanks to Prof. Dr. Hatice SOFU, above all, for the initial encouragement she gave me in pursuing my interest in linguistics research dating back to my undergraduate level. Her professional attire in academy and motivation for international projects shaped my ideas and views universally.

I would like to thank Prof. Dr. Hatice SOFU and Prof. Dr. Tahir BALCI for accepting to be members of examining committee for my thesis and also for their constructive feedback and encouraging attitude.

I would like to express my heartfelt gratitude to Asst. Prof. Dr. Mehmet SEYİS, Asst. Prof. Dr. Rana YILDIRIM and Asst. Prof. Dr. Hasan BEDİR for the suggestions they have given in the analysis and interpretation of the data.

I owe special thanks to my precious friends Betül UÇAR and Abdullah PEKEL for their marvelous support and understanding. Without their support and aspirations, I would have had a lot of problems. Also I would like to send my very special thanks to my MA classmates especially my friend Yunus Emre AKBANA.

My family is the invisible force who has made this thesis successful and possible today. My eternal thanks go to my beloved father and mother, who both have inspired me to learn and search more in life.

Lastly, I owe special thanks to all people who contributed to my study.

TABLE OF CONTENTS

Page

ÖZET ..... **iii**

ABSTRACT ..... **v**

ACKNOWLEDGMENTS ..... **vii**

LIST OF ABBREVIATIONS ..... **xi**

LIST OF TABLES ..... **xii**

LIST OF FIGURES ..... **xiv**

CHAPTER 1

INTRODUCTION

1.1. Background of the Study 1

1.2. Corpus Linguistics 3

1.3. The History of Corpus Linguistics 4

1.3.1. Pre-Corpus Linguistics 6

1.3.2. Rationalist and Empiricist Views on Corpus Linguistics 7

1.3.3. Modern Corpus Linguistics and Corpus Based Studies 9

1.4. Statement of the Problem 10

1.5. Aim of the Study 10

1.6. Research Questions 11

1.7. The Significance of the Study 11

1.8. Limitations of the Study 11

CHAPTER 2

REVIEW OF LITERATURE

2.1. Introduction 13

2.2. Interlanguage 13

2.2.1. Contrastive Interlanguage Analysis 14

2.3. Qualitative and Quantitative Analyses 15

2.4. Learner Corpus 16

2.5. Data Driven Learning 18

2.6. Corpus Based Researches on Relative Clauses with the Relative Pronoun *Which* 20
2.7. Relative Clauses 22
2.7.1. English Relative Clauses 22
2.7.1.1. Restrictive and Non-Restrictive Relative Clauses in English 22
2.7.2. Turkish Relative Clauses 23
2.7.2.1. Restrictive and Non-Restrictive Relative Clauses in Turkish 25
2.7.3. Japanese Relative Clauses 25

CHAPTER 3
METHODOLOGY

3.1. Introduction 27
3.2. Research Design 27
3.3. Materials 27
3.3.1. Corpora Used in the Study 28
3.3.2. ICLE 29
3.3.2.1. Genre 30
3.3.2.2. Topic 32
3.3.2.3. Task Setting 33
3.3.2.4. Learner Variables 33
3.3.2.4.1. Age 33
3.3.2.4.2. Gender 34
3.3.2.4.3. Mother tongue 35
3.3.2.4.4. Region 35
3.3.2.4.5. Knowledge of other foreign languages 35
3.3.2.4.6. Time Spent in an English-Speaking Country 36
3.3.2.4.7. Learning Context 36
3.3.2.4.8. Proficiency 36
3.3.3. LOCNESS 38
3.3.4. TUC 38
3.3.5. Software 38
3.4. Data Analysis 40

CHAPTER 4
RESULTS AND DISCUSSION

4.1. Introduction 41
4.2. Analysis of the Data from Corpora 41
4.3. Analyses of Errors in TICLE learner corpus 49
4.3.1. Errors in the relative pronoun *which* in Restrictive Relative Clauses 49
4.3.1.1. Concord errors in the relative pronoun *which* in Restrictive Relative Clauses 49
4.3.2. Errors in the relative pronoun *which* in Non-restrictive Relative Clauses ... 52
4.3.2.1. Punctuation errors in the relative pronoun which in Non-restrictive Relative Clauses 52
4.3.2.2. Errors of Sentential Relative Clauses 54
4.3.3. Choice of Relativizer Pronoun 55
4.4. Analyses of Errors in JPICLE learner corpus 55
4.4.1. Errors of the relative pronoun *which* in Restrictive Relative Clauses 56
4.4.1.1. Concord errors in the relative pronoun *which* in Restrictive Relative Clauses 56
4.4.2. Errors of the relative pronoun *which* in Non-restrictive Relative Clauses ... 57
4.4.2.1. Punctuation errors in the relative pronoun which in Non-restrictive Relative Clauses 57
4.4.2.2. Errors of Sentential Relative Clauses 58
4.4.3. Choice of Relativizer Pronoun 59

CHAPTER 5
CONCLUSION

5.1. Introduction 60
5.2. Evaluation of Research Questions 60
5.3. Implications for English Language Teaching 62
5.4. Further Research 63

REFERENCES 64
CURRICULUM VITAE 68

LIST OF ABBREVIATIONS

L1 = First Language

L2 = Second Language

ELT = English Language Teaching

ICLE = International Corpus of Learner English

TICLE = Turkish International Corpus of Learner English

JPICLE = Japanese International Corpus of Learner English

TUC = Turkish University Corpus

LOCNESS = Louvain Corpus of Native English Essays

LLC = Longman Learners' Corpus

HKUST = Hong Kong University of Science and Technology

CA = Contrastive Analysis

CIA = Contrastive Interlanguage Analysis

IL = Interlanguage

DDL = Data Driven Learning

NL = Native Language

RC = Relative Clause

NP = Noun Phrase

VP = Verb Phrase

KWIC = Key Word in Context

LIST OF TABLES

Page

Table 1. Corpus Linguistics from 1950s to 1980s 5
Table 2. DDL material aiming to set up learner-centered corpus analysis..... 19
Table 3. Summary of Strategies of Relativization in Turkish 24
Table 4. The size of the corpora..... 28
Table 5. ICLE task and Learner Variables 29
Table 6. Proportion of Argumentative Essays 30
Table 7. Average Essay Length 31
Table 8. Top ten essay topics..... 32
Table 9. Learners' Age in JPICLE and TICLE..... 33
Table 10. Learners' gender 34
Table 11. Languages spoken in more than one country 35
Table 12. CEF Results – 20 essays per subcorpus..... 37
Table 13. The overall use of the relative clauses with the relative pronoun *which* in TICLE, JPICLE, LOCNESS and TUC 41
Table 14. The Comparison among Turkish and Japanese learner corpora and the native corpora in the relative clauses with the relative pronoun *which*..... 42
Table 15. The overall use of the restrictive relative clauses with the relative pronoun *which* in TICLE, JPICLE, LOCNESS and TUC..... 43
Table 16. The Comparison among Turkish and Japanese learner corpora and the native corpora in the restrictive relative clauses with the relative pronoun *which* 44
Table 17. The overall use of non- restrictive relative clauses with the relative pronoun *which* in TICLE, JPICLE, LOCNESS and TUC..... 45
Table 18. The Comparison among Turkish and Japanese learner corpora and the native corpora in the non-restrictive relative clauses with the relative pronoun *which* 46
Table 19. The overall errors in the relative clauses with the relative pronoun *which* in TICLE and JPICLE..... 46

Table 20. The overall errors in the restrictive relative clauses with the relative pronoun *which* in TICLE and JPICLE..... 47

Table 21. The Comparison of errors among Turkish and Japanese learner corpora and the native corpora in the restrictive relative clauses with the relative pronoun *which* 47

Table 22. The overall errors in the non-restrictive relative clauses with the relative pronoun *which* in TICLE and JPICLE..... 48

Table 23. The Comparison of errors among Turkish and Japanese learner corpora and the native corpora in the non-restrictive relative clauses with the relative pronoun *which* 48

LIST OF FIGURES

Page

Figure 1. A screenshot of the relative pronoun *which* (from LOCNESS) by WordSmith..... 39

Figure 2. A screenshot of the relative pronoun *which* (from TICLE) by ICLE 39

CHAPTER 1

INTRODUCTION

This section consists of background of the study, the overview of corpus linguistics which constitutes the basis of the study, the statement of the problem, aim of the study and research questions.

1.1. Background of the Study

Wherever and whenever human beings exist, language exists. When buying something, playing a game or travelling on a public bus, communication plays a crucial role. Language has always been a lively tool as being the mirror of each era in humanity for centuries. We learn about our past via language revealing the secrets of that specific period. Since all languages change through time, we have always been interested in every single detail of any language for understanding the languages and their relationship with others. In this sense, society and mind are inextricable for creating a language. Although the language itself is broader than an individual, each individual actually creates a unique language product. For Saussure (1966, p. 23) ‘‘Language is a social product deposited in the brain of each individual.’’

Linguistics, which is a field of study composed of many different pursuits, researching sounds, grammar, language learning, and language history, encompasses a number of subfields such as semantics, morphology, syntax, sociolinguistics, language acquisition, psycholinguistics and so forth. In the history of language pedagogy, better ways, methods or tips for teaching and learning another language have been the main focuses of linguistics’ subfields due to the increasing demand of globalization. Specifically, the research on methods and approaches on how to teach and learn English language has attracted more attention than the ones for any other language. Corpus linguistics, which is regarded as both a subfield of linguistics and a methodology, has been in growth in recent years as to language learning and teaching researches. Although considered as a young trend, the studies on corpus linguistics have a very long and lively history dating back to earlier eras from slip cards to modern and fast computers. Corpus linguistics does not merely mean extracting language data through the use of computers. On the contrary, corpus linguistics is the study and analysis of

data extracted from a corpus. The corpus linguist is not there to find the data but to analyze it by means of computers (McEnery & Wilson, 2001, pp. 1-12).

McCarthy and O'Keeffe (2010, p. 3) state that corpus linguistics is generally associated with the linguists searching through screen after screen of concordance lines and wordlists generated by computer software in order to make the language items more descriptive in big texts or big collections of smaller texts. As a matter of fact this method based on detailed searches for words and phrases in the context dates back to the thirteenth century.

According to McEnery and Wilson (2001, p. 1) ''Corpus linguistics is perhaps best described for the moment in simple terms as the study of language based on examples of 'real life' language use.'' In recent years researches on real life language use of learner corpora have been increasing day by day. Real language data for a course book could be very useful as showing the errors students make and highlighting the parts which cause problems for students. On the one hand, the learner corpus data could help the ELT teachers predict the possible errors of students beforehand and prepare specific-to-target activities to eliminate errors. A number of major English course book publishers have started to base their content on international learner corpus. For instance, *Cambridge face2face* states the motto 'real English guarantee' by referencing Cambridge International Corpus, which is a collection of 800 million words of real spoken and written English including 60,000 exam papers from *Cambridge ESOL*. In essence, corpus linguistics is an interdisciplinary field which cooperates with the other subfields of linguistics such as linguistic theory, second language acquisition, foreign language teaching. Furthermore, it has been used as an aid tool for grammar books, course books, and dictionary databases. In this perspective, Granger (1998, p. 5) explains that any ELT teacher's interest in corpus linguistics could be connected to improving her own teaching methods and practices. For instance, she will be able to save time where the students experience no difficulty and focus on areas of language items where more help and practice is needed. James (1992, p. 190) also states that ''the really authentic texts for foreign language learning are not those produced by native speakers for native speakers, but those produced by learners themselves.''

In this current study we aim to describe the use and errors of relative pronoun *which* in Turkish university students' written English by comparing it with LOCNESS, TUC and JPICLE.

1.2. Corpus Linguistics

In recent years, the importance given to the language learning has reached a level at which everybody wants to be as accurate and fluent as possible especially in the production of English, as it is the lingua franca of the world. To seek better and alternatively effective ways for teaching and learning English has given a momentum to corpus linguistics today. As a matter of fact, the studies on corpora are not new. On the contrary, it dates back to the 1350s with Amarna letters, which constituted the diplomatic archive of correspondence clay tablets between the Egyptian administrations (http://en.wikipedia.org/wiki/Text_corpus).

For many centuries, a great many various types of studies on corpus linguistics have been carried out. For instance, diary studies of child language by Preyer in 1889 and by Stern in 1924, research into language pedagogy by Fries in 1940 and Traver and by Bongers in 1947, spelling conventions by Kading in 1897 and lastly comparative linguistics by Eaton in 1940 used large bodies of text and provided the bridge to the period of 1960s. The Survey of English Usage (SEU) in 1960 by Quirk and later Brown Corpus in 1964 by Francis and Kucera can be called as a real milestone of the modern corpus linguistics. In fact, corpus linguistics started to grow and be accepted as a valid means of language enquiry (Baker, Hardie & McEnery, 2006, pp.50-51).

The confusion over whether corpus linguistics is merely a methodology rather than a linguistic field of study does not prevent us from placing the corpus linguistics in the heart of nearly all linguistic fields. McEnery and Wilson (2001, p. 2) state that ‘‘Corpus linguistics is a methodology that may be used in almost any area of linguistics, but it does not truly delimit an area of linguistics itself.’’ Therefore, corpus linguistics can include the studies from various fields such as lexicology, semantic studies, grammatical studies, sociolinguistics, and interlanguage studies.

Kennedy (1998, p. 1) states that ‘‘corpus linguistics is not an end itself but is one source of evidence for improving descriptions of the structure and use of languages, and for various applications, including the processing of natural languages by machine and understanding how to learn and teach a language.’’

As regards to the definition of what the corpus linguist is, Chafe (1992, p. 36) states that ‘‘the corpus linguist is a linguist who tries to understand language by carefully observing extensive natural samples of it and then, with insight and

imagination, constructing plausible understandings that encompass and explain those observations''.

1.3. The History of Corpus Linguistics

The history of Corpus linguistics encompasses three main stages in chronological order: Pre-corpus linguistics, the period between 1950s and 1980s and the modern corpus linguistics.

The first stage is pre-corpus linguistics or early corpus linguistics in which the studies were mainly based on traditional methods of collecting data involving priests, monks or university staff in the collection and analysis process. Meyer (2004, p.11) stresses the relationship of pre-corpus linguistics and its methodology by giving an example:

> Pre-electronic corpora were created prior to the computer era, consisted of a text or texts that served as the basis of a particular project, and had to be analyzed through often time-consuming and tedious manual analysis. For instance, in the 18th century, Alexander Cruden created a concordance of the King James Version of the Bible, an extremely work-intensive effort that had to be done with pen and paper.

According to McEnery and Wilson (2001, p. 2) early corpus linguistics is a term which is used to describe linguistics before the advent of Chomsky. The discussion between rationalists especially Chomsky and the empiricists aided corpus linguistics to be understood better. Therefore, we will mention the rationalist and empiricist views on corpus phenomena in another subtitle in order to make the breaking point clearer. The corpus linguistics period between 1950s and 1980s, which seemed to be in a discontinuous process, was another important stage as to the development of modern corpora today as the pioneers such as Roberto Busa, Jan Svartvik, Geoffrey Leech, J. R. Firth and John Sinclair carried out researches on the corpus data. In this second important stage, the works by humanities computing, mechanolinguistics, the study of English grammar and neo-Firthian were outstanding as to reviving the corpus studies and connecting to the modern era.

Table 1

Corpus Linguistics from 1950s to 1980s

Computing corpus linguistics	
Pioneer(s):	Roberto Busa
Focus of the study:	Earlier computational corpus-based linguistics
Period:	1947-1964
Name of the corpus project:	The first study collaboration with IBM computers
Machine readable corpus linguistics	
Pioneer(s):	Alphonso Juiland
Focus of the study:	Sampling techniques, text processing
Period:	1956s
Name of the corpus project:	Multilingual machine readable corpus
Study on English Grammar	
Pioneer(s):	Quirk, Svartvik, Francis and Kucera
Focus of the study:	Use of English grammar
Period:	1960-1980
Name of the corpus project:	Survey of English Usage (SEU), International Corpus of English, Brown Corpus London-Oslo Bergen (LOB)
Neo-Firthians Corpus Linguistics	
Pioneer(s):	J. R. Firth, John Sinclair
Focus of the study:	The examination of complete texts, an open-corpora
Period:	1950s
Name of the corpus project:	Cobuild Project, Bank of English

In the last stage 1980s and today, modern corpus linguistics has started to get more attention by linguists and scholars pursuing various fields of language study.

Meyer (2004, p. 8) states that ‘‘Electronic corpora are the mainstay of the modern era and a consequence of the computer revolution, beginning with the first computer corpora in the 1960s, such as Brown Corpus (Kucera/Francis 1967) and continuing to the present time.’’ By means of computer corpus based data, analyzing a huge amount of data to be used in any linguistic field is considerably easier and faster today.

1.3.1. Pre-Corpus Linguistics

Meyer (2004) summarizes the important phases of pre-corpus linguistics era in four parts, which were Biblical Concordances, Grammars, Dictionaries and The Survey of English Usage (SEU) Corpus. If we examine the phases of pre-electronic corpora, we firstly need to start with Biblical concordances representing the first examples of corpus based research with linguistic associations. They were written in many languages such as Latin, Greek, Hebrew, and English. Secondly, early grammars dealt with the use of corpora in the compilation of grammars. They compiled the grammars in that they could base their analyses on examples taken from the prominent writers of English such as early dictionaries, corpora have a long tradition in lexicography since they provide a source for illustrative quotations and they help the researchers receive the meaning of a word from the context. The fourth phase of pre-electronic corpora was the Survey of English Usage (SEU) Corpus, which was a corpus whose compilation began in 1959 at the Survey of English Usage (University College London) under the direction of Randolph Quirk. The SEU Corpus occupies an important place in the development of modern corpus linguistics itself due to being the first corpus created clearly for use (p.1-20). Comprehending the reason behind to create such a corpus is as crucial as the corpus itself. In fact, The SEU Corpus has an important place in the development of modern corpus linguistics as it was the first corpus to aim at practical use. Quirk (1974, p. 167) states that ‘‘the big grammars fell short for two main reasons. They were based on writing, not speech and their eclectic use of written source of materials too often leaves unclear the distinction between normal and relatively abnormal structures and the conditions for selecting the latter.’’

Even though the pre-electronic corpora seemed obsolete and unnecessary when compared the contemporary examples, those corpora had important influences on the development of corpus linguistics by creating a basis for language studies. In short, pre-

electronic corpus linguistics has a rich history that definitely contributed to the current situation of discipline.

1.3.2. Rationalist and Empiricist Views on Corpus Linguistics

Throughout the long and lively history of corpus linguistics, corpus linguists and other linguists have been discussing and sharing their ideas with each other. As regards to the rationalist and empiricist views on corpus linguistics, Chomsky's (1984) critics on corpus linguistics occupy fundamental points of views forming the generative grammarians' views towards corpus linguistics. For instance, according to Chomsky's (1984) criticisms, a linguist never needs a corpus since a linguist must model language as competence rather than performance. Moreover, a corpus is by its very nature a collection of externalized utterances; it is performance data and it must of necessity be a poor guide to modeling linguistic competence. It is competence rather than the performance that the linguist is trying to model as performance is just a poor mirror of competence. Rationalist view describes competence as being our tacit, internalized knowledge of a language and performance as being external evidence of language competence and its usage on particular situations when, especially, other factors may affect its form rather than competence. In short, rationalists supported artificially induced observations whereas empiricists supported naturally occurring observations. Criticisms by rationalist view shed light onto many foresights that directed the current modern corpus into an effective field of linguistics (McEnery & Wilson, 2001, p.11).

Meyer (2004, p. 10) also focuses on the perspectives of a corpus linguists and a generative grammarian. In corpus linguistics, the main issue is concerned about descriptive aspect rather than explanatory aspect. That is to say, while a corpus linguist aims to describe a language, a generative grammarian aims to focus on the theoretical aspect of a language by the knowledge of universal grammar. Criticisms on corpus data received a reaction from linguists who wanted to use corpus data. The effect of criticism on early corpus linguistics in the 1950s was hasty and in an extreme way. Chomsky's views were very influential indeed and in part valid. Those hasty and extreme criticisms did not make the corpus totally obsolete or unreliable tool, on the contrary, responding his criticisms helped the development of corpus research and also these criticisms have helped to foster a more realistic attitude towards corpora. As we have discussed in previous paragraphs, harsh criticisms of corpus-based approach were about the nature of

data, which was viewed as a poor mirror of competence. Chomsky (1984) criticized harshly naturally occurring data as 'degenerate quality' and added that collecting quantitative information is a waste of time and of no use to linguists. On the other hand, the artificial aspect of data in introspection based judgments seems one of the weakest links. As the researcher can influence the informant, s/he can artificially influence the data itself. This situation causes the informant to give typical introspective judgment. As a matter of fact, the introspection process might not be systematic and it focuses on the extremes reached by non-corpus linguist (McEnery & Wilson, 2001, p. 20).

As we have discussed before, in the lively history of corpus linguistics there have been a great many satires, harsh criticisms. For instance, Fillmore (1992, p. 32) states the relationship of a non-corpus linguist and a corpus linguist in a humorous way:

> Armchair linguistics does not have a good name in some linguistics circles. A caricature of the armchair linguist is something like this. He sits in a deep soft comfortable chair, with his eyes closed and his hands clasped behind his head. Once in a while he opens his eyes, sits up abruptly shouting, 'Wow, what a neat fact!' grabs his pencil, and writes something down. Then he paces around for a few hours in the excitement of having come to still closer to knowing what language is really like.

McEnery and Wilson (2001, p. 12) explain that the criticisms on corpus linguistics provided a well-established standing point and helped the field develop. On the other hand, a corpus aims to model performance instead of competence. Counting everything in any language one by one or describing a language fully is neither achievable nor plausible since natural languages are not finite. As for the melding of corpus and intuition, McEnery and Wilson (2001, p. 23) also state that ''the corpus and the linguist's intuition were complementary, not antagonistic.''

Stubbs (2001, p. 143) defines that Chomskyan linguists are interested in what speakers can say, and corpus linguists look at what they say. However, there can be a melding of the approaches. The findings from the corpus are not an interpretation of the meaning of the text, but a presentation of some of its formal features which the human reader can then more easily interpret. Fillmore (1992, p. 35) mentions about the usefulness of collaboration between corpus and non-corpus linguists as regards to describing and explaining language items:

> I do not think there can be any corpora, however large, that contain information about all of the areas of English lexicon and grammar that I want to explore, but every corpus I have had the chance to examine, however small, has taught me facts I could not imagine finding out any other way. My conclusion is that the two types of linguists need one another.

In the history of corpus linguistics, the exchange of views among linguists has proven the field of corpus linguistics as a mediator that can establish a bridge between introspection and performance.

1.3.3. Modern Corpus Linguistics and Corpus-Based Studies

In the frame of modern corpus linguistics, the most prominent feature of corpus linguistics is that it is regarded as a methodology rather than a linguistic area of study. The prevalent use of computer corpus has emerged from the corpus linguistics as a new discipline and has already provided the researchers utilize such big collections of corpora as the International Corpus of Learner English (ICLE), which is a corpus of learner English from several mother tongue backgrounds, the Longman Learners' Corpus (LLC) and Hong Kong University of Science and Technology Learner Corpus (HKUST).

Baker et al. (2006) claim that today the most common types of corpus - based studies are gathered around learner corpus studies generally based on grammatical structures which are frequently carried out at schools and academic surveys. Baker et al. (2006, p. 103) also state that a learner corpus can be regarded as the output obtained from learners and the compilation of written essays using pre-set topics produced in language teaching classrooms. For instance, Biber (2006) researched on grammar features that are commonly used in university spoken and written registers by including both instructional registers and student advising/management registers such as office hours, class management talk, written syllabi. As regards to providing more examples of grammatical studies based on corpus, Schmied's (1993) study of relative clauses provides quantitative knowledge about aspects of relative clauses in the LOB and Kolhapur corpora. Oostdijk and De Haan's (1994) study focused on analyzing the frequency of various English clause types in the Nijmegen Corpus. Furthermore,

Halliday's (1991) theory systemic grammar, which was based on probabilities of language systems, aimed to differentiate the probabilities of relativizer *which* and *that* (McEnery & Wilson, 2001, p. 109).

1.4. Statement of the Problem

It is a fact that the main goal of studies conducted in SLA field is to reveal the principles that rule the second language learning process. However, the process of learning a second language is mental and many of the studies based on introspective data and thus it is not directly observable (Granger, 1998). In fact, a great number of SLA studies, which depend on elicited data causing learners to produce a language differing from real language, also could bring limitation to the generalization of the results. Thus, a corpus based study, depending on the learner performance data, can provide a clear and a concrete frame in terms of representing language items. In recent years, the number of studies conducted in corpus linguistics has increased substantially. Yet, the same is not true of the research made on the corpus based study of relative clauses *which* pronoun in Turkish students' academic texts. For this very reason, the present study is concerned with the use of relative clauses with the relative pronoun *which* in Turkish university students' argumentative essays in English. It particularly describes the use of relative pronoun *which* both in restrictive and non-restrictive clauses. As a matter of fact, the relative pronoun *which* is considerably common in written texts since it can specify the reference of the head noun in restrictive relative clauses and give elaborating information in non-restrictive relative clauses. In this context, it is expected that the findings of this study will provide the differences in the relative pronoun *which* usage between non-native speakers (NNS), namely Turkish and Japanese and native speaker (NS), American English Native Speakers in terms of overuse or underuse. Thus, this study intends to suggest a descriptive frame as regards to overuse, underuse frequency and error types.

1.5. Aim of the Study

The main purpose of this study is to examine the use of relative clauses with the relative pronoun *which* in Turkish university students' written English as to underuse, overuse and error types and compare it with the use in three aspects: native English

writing, native Turkish writing and learner data Japanese university students' written English. The study also aims at shedding light into the difficulties of the relative pronoun *which* that students encounter in the process of dealing with written texts.

1.6. Research Questions

This study seeks answers to the following questions:

1. Do Turkish university students use the relative pronoun *which* both in restrictive and non-restrictive relative clauses to the same extent as native speaker American university students?
2. Does Turkish university students' use of the relative pronoun *which* both in restrictive and non-restrictive relative clauses give evidence of transfer from their native language Turkish?
3. Do relative clauses with the relative pronoun *which* occur most frequently as restrictive or non-restrictive clauses in the texts?
4. If there is a significant difference in terms of underuse or overuse of relative pronoun *'which'*, does this stem from the Turkish learners' L1 or is it just the interlanguage universal?

1.7. The Significance of the Study

Not many studies have been conducted on the corpus-based data of relative clauses with the relative pronoun *which.* Therefore, this present study focuses on the use of relative clauses with the relative pronoun *which* in Turkish university students' written English as to underuse, overuse and error types and is to comparison on the use in four corpora: native English writing, native Turkish writing, Turkish learners' L2 English writing and the learner data from Japanese university students' written English.

1.8. Limitations of the Study

This present study is only limited to the learner corpus ICLE including two subcorpora TICLE and JPICLE and the two native corpora LOCNESS and TUC. In terms of comparing with a learner corpus, it is limited to only one learner language data

JPICLE. Furthermore, the focus of the study is limited to the relative pronoun *which* rather than including the other relative pronouns such as *that*, *who*, *whose* etc.

CHAPTER 2

REVIEW OF LITERATURE

2.1. Introduction

This chapter intends to present the relevant literature for the study. In this context, Interlanguage, Contrastive Interlanguage Analysis (CIA), Qualitative and Quantitative Analyses, Learner Corpus, and Data Driven Learning (DDL) have been presented. The relative clauses with the relative pronoun *which* have been presented in the frame of previous studies and descriptive grammar.

2.2. Interlanguage

In second language acquisition research, the underlying knowledge is that L2 learners construct their own unique, creative, and transformational linguistic system when they start to acquire a second language. In fact, this consistently changing linguistic system is known as an interlanguage, theoretically introduced by Selinker (1972). He provided the concept of developing second language knowledge. Lightbown and Spada (1999, p. 77) state that the nature of IL is systematic, dynamic and influenced by the learners previous acquired languages. Learners can continually apply to their IL by adding rules, deleting or reconstructing rules, creating new strategies and they reach results as interlanguage continuum. Furthermore, each IL is an idiosyncratic variety of the target language created by a particular learner and each IL is unique, the incomplete product of the L2 learning process, not an existing variety available to be learned.

Selinker (1972, p. 219) explains five main characteristics of an interlanguage process:

1. *First language transfer* : Learners might transfer from the native language mostly in the form of interference.
2. *Transfer of training:* Learners might be exposed to particular methods and techniques used in training and this might cause the learners adopt the most frequently repeated item by the instructor.

3. *Strategies of second language learning:* This characteristic is related to the strategies to the language items being learned and the learners tend to simplify the items so that they can reduce the target language to a simpler system.
4. *Strategies of second-language communication:* Learner's strategies to manage communication in the target language by means of avoiding problematic items.
5. *Overgeneralization of target language linguistic material*: Learners tend to overgeneralize of target language rules such as adding –ed ending on an irregular verb 'catched' instead of 'caught'.

IL, then, as stated by Selinker (1972, p. 214), is ''the separate linguistic system based on observable output which results from a learners' attempted production of a target language norm,'' where the target language is defined as ''one norm of one dialect.'' Thus, the language L2 learner comes up with is not the actual L2. It is only ''an attempted production'' of L2. It neither completely shows the features of L1 called as language transfer nor does it completely represent the characteristics of the L2. As a matter of fact, IL is a unique, systematic and dynamic linguistic system.

2.2.1 Contrastive Interlanguage Analysis

Contrastive Analysis (CA) compares different languages. Lado (1957) states the main assumption of CA is that ''individuals tend to transfer the forms and meanings and the distribution of the forms and meanings of their native language'' (p. 2). Lightbown and Spada (1999, p. 78) also state that CA is based on the premise that errors were generally assumed to be the result of transfer from learners' first language. Not all errors made by second language learners can be explained in terms of the first language transfer alone. All errors made by learners could not be explained by CA. Thus, a detailed description and analysis of errors was necessary (Lightbown & Spada, 1999, p. 79). However, Contrastive Interlanguage Analysis (CIA) involves the comparison and contrasting what non-native and native speakers of a language do in a comparable situation. In essence, CIA involves two major types of comparison: comparison of native language and interlanguage NL vs. IL or comparison of different languages IL vs. IL. For instance, native language and interlanguage comparisons aim to uncover the

features of non-nativeness of learner language. These comparisons contribute to SLA theory by providing insight into the role of transfer and they help the teachers to reach their targets on the basis of learner needs (Granger, 1998, p. 17).

Most of the studies based on ICLE database have focused on quantitative differences between L1 and L2 instead of on L2 errors. Thus, the main aim of contrastive interlanguage analysis is to examine differences in frequency between L1 and L2. Furthermore, these comparisons of L1 and L2 help us to evaluate whether the differences are L1 dependent or due to their mother tongue background. When a linguist would like to conduct a research based on CIA, s/he is supposed to focus on one specific linguistic enquiry such as modals, relative clauses or sentence connectors. After choosing this linguistic area of study, s/he needs software such as *Wordsmith tools (WST)* in that s/he can extract the concordance lines. By using concordance, it is quite easy to retrieve the words that are significantly over- or underused by learners. It is crucial to ensure comparability of the L2 data and native English data. Therefore, the *Louvain Corpus of Native English Essays (LOCNESS)* has been used as a reference corpus to compare the *ICLE* data. If another type of native English corpus such as the LOB corpus rather than the *LOCNESS* was used to compare with the ICLE data, the results of the analysis would be quite different, for the collection and design criteria of the *ICLE* and *LOCNESS* are quite similar. For this very reason, the corpus based studies using contrastive interlanguage analysis generally depend on L1-L2 comparisons based on the *LOCNESS* corpus (Granger, Dagneaux, Meunier & Paquot, 2009, p. 43). Therefore, in this study we have based our study of the relative pronoun *which* in Turkish university students' written English on Contrastive Interlanguage Analysis (CIA).

2.3. Qualitative and Quantitative Analyses

As we have discussed in the previous part, with the help of computing technology especially the software for concordance analysis, the research on corpus has become much faster and more reliable. Biber (1998, p. 4) states that qualitative analysis is crucial since it provides the interpretation of the pragmatic role of metaphors. On the other hand, quantitative analysis is crucial because it provides the basis for judging norms of language use. To understand the difference and cooperation between two analyses, Charteris (2004, p. 32) states that:

> Quantitative approaches are concerned with frequency and typicality – to answer questions such as how common is a particular word or phrase in the language is general. Or how typical is particular form used within one section of the corpus as compared with another. They are quicker to conduct because they are not concerned with the context of the target forms. Qualitative approaches answer questions such as: what are the different meanings that are attached to particular words or phrases? Are senses literal or metaphoric? What type of evaluations do they convey?

In this study, we aim to examine the quantitative data obtained from the learner and native corpora. The main purpose for using the quantitative analysis is to investigate the frequencies of the usages of relative pronoun *which* both in restrictive and non-restrictive relative clauses. As a result of this, we would like to answer our research questions about frequency such as which type of relative clause is more common or less common, how they are used in the texts and why they are used so. When a specific linguistic enquiry is used frequently by learners, it could give us more momentum to search about it.

2.4. Learner Corpus

There are a great many corpora used in corpus linguistics research such as specialized, reference, multilingual, parallel, diachronic, monitor and learner corpora. These corpora have their own purpose. For instance, a specialized corpus is used in particular research projects such as giving examples about business English (Baker et al. 2006, p. 49). However, in this study we have focused on leaner corpus as it is a type of corpus obtained from learners. Learner corpora consist of written essays generally on the topic of argumentative essays (Baker et al. 2006, p. 103).

The learner corpus has filled an important gap for language teaching and learning. Most of the course books such as *Cambridge Face to Face* or *English Unlimited* have started to utilize learner corpora in designing the syllabus. In this context, James (1992, p. 190) states that ‘‘the really authentic texts for foreign language learning and teaching are not those produced by native speakers for native speakers, but those produced by learners themselves’’. As regards to mentioning about authentic or non-authentic English in the frame of language teaching, Beaugrande (2001, p. 23)

claims that ‘‘some non-native teachers of EFL suffer partly from exposure to non-authentic English and partly from lack of exposure to authentic Standard English’’. Granger (2002, p. 7) also states that learner corpora are collections of authentic textual data assembled according to explicit design criteria and they can be used for a particular foreign language teaching purpose. Nessalhauf (2004, p. 9) also claims that leaner corpora involve in authentic texts to make use of because learner corpora can provide a basis for the description of frequently occurring mistakes in language learning process. Therefore, learner corpora can be regarded as more natural than any type of corpora since it is the collection of texts from non-native speakers of the target language.

Biber and Reppen (1998) state that ‘‘learner corpora use in SLA research is crucial in terms of examining frequency and persistence of errors in second language learner groups. Such studies increase our understanding of second language acquisition, provide data for other perspectives on errors and provide evidence for instructional decisions.’’ Tono (2003) summarizes some examples of researches based on learner corpus data which include topics adverbial connectors (Altenberg and Tapper 1998; Lorenz 1998), multiword units (De Cock 1998), direct questions (Virtanen 1998a), the progressive (Virtanen 1998b), tense morphology (Granger 1998) and phrasal verbs (Lam & Hung 1998), an analysis of complement clauses (Biber & Reppen 1998). To give a broader example of a learner corpus based study, Lorenz (1999) researches on adjective intensification in Learner English by comparing the German sub corpus of ICLE and comparable native corpus. After completing the study, he reaches the results below:

- Overuse: German learners of English use linguistic forms significantly more frequently than native speakers of English. For instance, the item *‘really’* is overused such as ‘*really big, really important, and really interesting’*.
- Underuse: German learners of English use linguistic forms significantly less frequently than native speakers of English. For instance, the item *‘particularly’* is overused such as ‘*particularly difficult, particularly useful’*.
- Misuse: German learners of English use the forms wrongly due to the interference. For instance, these learners tend to add ‘*maximizer’* to non-gradable adjectives such as ‘*absolutely silly, totally damaged*’. Furthermore, they add forms that do not exist in English such as ‘*a too complex problem’*.

As a matter of fact, the analysis of learner corpora and its comparison with comparable native corpora provide important quantitative and qualitative insights into learner's knowledge of English whether s/he approached to or deviate from the target language model.

2.5. Data Driven Learning (DDL)

As we have discussed in the previous chapters, the field of corpus linguistics is considerably extensive from lexicography to ELT classrooms. We have the chance of utilizing quite modern and useful corpus-based dictionaries such as Collins *COBUILD English Language Dictionary* (1987) and *Macmillan English Dictionary* (2002). While corpus-based dictionaries include the frequency and grammatical description of relevant word, corpus-based grammars fulfill the grammatical description of specific syntactic structures. For instance, The *Collins COBUILD English Grammar* (1990) is a learner grammar and semantically-oriented (Mukherjee, 2004, p. 6). As regards to using corpora in the ELT classrooms, there have been a great many suggestions by corpus linguists. Johns and King (1991) suggested the term *Data-driven Learning* in that they could label the learning via corpus as 'learning as discovery' and 'learner-centered autonomous approach:

Data-driven learning is the use of computer generated concordances in the classroom to get students to explore regularities of patterning in the target language or the development of activities and exercises based on concordance output (John and King, 1991, p. iii). Mukherjee (2004) explains the pedagogical aspects of corpus-based activities into three categories: for dictionaries and other materials for ELT classroom, as a database and tool in the ELT classroom itself, and as representative samples of learner language. The DDL material in Table 2 exemplifies concordance-based exercises:

Table 2

DDL Material Aiming To Set Up Learner-Centered Corpus Analysis (Gavioli, 2001, p. 120)

---

handsome good attractive pretty tall successful

1. In pairs, choose three of these adjectives and have a look at their concordances.
    - Which words do they most frequently occur with?
    - Do they occur as part of a series of adjectives?
2. Discuss your findings with other pairs, then with the rest of the class.

By doing the particular exercise in table 2, students can be more conscious of frequencies in the target language and contextual use in the sentences. As a matter of fact, the DDL activities can be regarded as learner autonomy, ranging from teacher-led and relatively closed concordance-based exercises to entirely learner-centered corpus projects (Mukherjee, 2004, p. 10). DDL has a number of advantages over a rule based approach. Firstly, many rules derived from intuition simply do not describe actual usage, as any number of corpus studies have shown over the last 20 years or more. Secondly rules and exceptions do not provide an accurate picture of language in general, which adheres to patterns, tendencies and generalizations of prototypical usage rather than rigid right or wrong. Thirdly, rules rarely give an idea of frequencies: one may teach beginners the use of perfect and continuous aspects, but forget to point out that 90% of all verb phrases are not marked for aspect. Rules tend to be considerably abstract because they aim to account for general language use. However, corpus data can show substantial differences in the use of grammar in registers, or text types both in speech and writing (Biber et al. 1999, p. 461). Although DDL might seem a revolutionary approach, it actually set a melting pot media between past activities and DDL activities. Accordingly, Gabrielatos (2005, p. 35) suggests that ‘‘the use of corpora should not be treated as an alternative to, or rival of existing teaching approaches, but as a welcome addition’’. Furthermore, Gavioli (2005, p. 40) claims that ‘‘DDL does not reject past practice, it builds on it, drawing on existing skills highly prized in the communicative classroom ’’. Mukherjee and Rohrbach (2006, p.20) summarize the pedagogical ends of using a learner corpus in the ELT classroom as presented below:

- By using corpus linguistic software like *WordSmith Tools*, the teacher can analyze the range of general and topic-related vocabulary that students in general or individual students have used.
- The learner corpus makes it possible to focus on individual learners' interlanguage and to provide them with tailor-made feedback.
- The learner corpus allows for longitudinal studies of learner language progression across time (both for the entire class and for individual learners). Thus, a local learner corpus has a monitoring quality.
- The corpus can be analyzed both by teachers and by students. For example, individual learners can use a concordance display of their own mistakes as a starting point for data-driven learning activities.
- The focus on the students' output will directly involve many more teachers in corpus-based activities and the exploration of learner data by the learners themselves will motivate many more learners to reflect on language use and thus raise foreign language awareness.

2.6. Corpus Based Researches on Relative Clauses with the Relative Pronoun *Which*

In recent years the corpus based research on relative clauses has been increasing as regards to both spoken and written data. There are two main types of clause in relative clauses as 'restrictive' and 'non-restrictive'. These two types of clause convey different types of information. Defining/restrictive information specifies something or someone, typically a noun phrase, by separating it from other members of a class. First of all, we are going to summarize some of researches based on spoken corpora on relative clauses. Cornilescu (1981) noted that restrictive relative clauses were typical after nouns modified by words such as any, no and every, while non-restrictive relative clauses typically occurred after proper names. Cornilescu (1981) also focused on the function of intonation pointing difference between restrictive and non-restrictive and came to conclusion that intonation provides the difference of these two types of relative clauses. Another example of corpus based research on relative clauses via utilizing the Lancaster/IBM Spoken English Corpus, Yamashita (1994) examined the positioning of

restrictive and non-restrictive clauses, and came to conclusion that non-restrictive relative clauses are more likely to occur in sentence-final position due to the fact that they convey rather more complex information than restrictive relative clauses. Depraetere (1995) also stated that non-restrictive relative clauses are more likely to convey specifying and particular information. Tao and McCarty (2001) investigated on the use of restrictive and non-restrictive relative clause in the frame of speaking and they classified the non-restrictive relative clauses into 'evaluative', 'expansion' and 'affirmative'. Each classification has their own function such as giving the speaker's opinion, attitude or stance, and giving additional information. That is to say, there is more flexibility in speaking rather than writing, written texts seem to have more restrictions such as punctuation, single authorization, and the researchers suggest that in terms of grammar pedagogy it is better to separate the written contexts from spoken contexts. In next paragraph we are going to focus on the corpus based research on relative clauses in the frame of written texts.

Most of the students who are studying at university meet a range of problems in the process of producing written texts in their L2 English. In fact, not only language learners of English but also native speakers of English need to adapt or aware of writing tasks in university language. Thereby, the importance of research on university students' written data has been increasing. Biber (2006, p. 2) states that there is more to carry out researches that document the important linguistic characteristic of academic registers in order to provide basis for corpus based grammar instruction. Furthermore, Biber (2006, p. 76) states that relative clauses have two primary functions: to specify the reference of the head noun, or to provide elaborating information. However, in many situations the relative clauses with the relative pronoun *which* serves both clause types 'restrictive' and 'non-restrictive'. Although the relative pronouns *which* and *that* are similar in their grammatical function, the relative pronoun *which* in academic texts is the more frequent than the latter. A more important difference is that *which* can be used in non-restrictive relative clauses. In contrast, *that* cannot be used in written academic texts despite its rare usage with non-restrictive clauses in fiction. The relative pronoun *which* has more conservative, academic associations and is thus used more in academic texts. According to the corpus findings from *Longman Spoken and Written English Corpus* (LSWE), there are more occurrences of restrictive relative clauses than non-restrictive relative clauses in native English texts of academic prose (Biber et al. 1999, p. 611). Thus, the findings and comparisons of various relative pronouns in the

LSWE corpus have inspired us to examine with the use in Turkish university students' written English. In this study, the relative pronoun *which* has specifically been chosen owing to the fact that it would give us the chance of researching both restrictive and non-restrictive relative clauses in written texts. Accordingly, we tried to describe the use of relative clauses with the relative pronoun *which* in Turkish university students' written English as to underuse, overuse and error types and compared it with its use in three aspects: native English writing, native Turkish writing and another learner data Japanese university students' written English. In addition to describing the use of a complex structure relative clauses with relative pronoun *which*, we tried to seek answers of our research questions on possible mother tongue interference or interlanguage feature of the structure.

2.7. Relative Clauses

2.7.1. English Relative Clauses

The main function of relative clauses in English is to modify a noun phrase. The term relative clause is used for various types of subclauses which are linked to the main clause. The linking is achieved with a back-pointing element, generally a relative pronoun. The relative pronouns are *who, whom, whose, which, that,* and *zero* pronoun.

1. The RECORDS ***which he owns*** are mostly classical.
[The relative pronoun ***which*** functions as object of ***owns***]

In the example 1, the relative clause is *which he owns.* The relative pronoun *which* points back to the head of the noun phrase (The RECORDS), which is called the **antecedent** as shown in small capitals (Leech & Svartvik, 2002, p. 383).

2.7.1.1. Restrictive and Non-Restrictive Relative Clauses in English

All the relative pronouns can be used in restrictive relative clauses. Among all these relative pronouns, relative pronoun *that* which is dependent of the personal on nonpersonal character of the antecedent and the *zero* relative are used more frequently than the other types of relative pronouns such as who, whom, which and whose

2. This author uses lots of WORDS ***which*** are new to me.

3. I need to talk about the E-MAIL ***which*** you sent me (Leech & Svartvik, 2002, p. 386).

As the focus of this particular study is to examine relative clauses both in restrictive and non-restrictive forms, the relative pronoun *which* has particularly been chosen. In constructing non-restrictive relative clauses only *wh*-pronouns are used. The meaning of a non-restrictive relative clause is quite similar to a coordinated clause with or without conjunction. Then non-restrictive relative clauses are shown separately in written language by punctuation (Greenbaum & Quirk, 2003, p. 370).

4. Here is Ronal Walker, *who (m)* I mentioned the other day.

5. He got lost on Snowdon, *which* was enveloped in fog.

One type of non-restrictive relative clause is sentential relative clauses. It is a special type of non-restrictive relative clause. It does not point back to a noun phrase but to a whole clause or sentence (Leech & Svartvik, 2002, p. 387).

6. The COUNTRY IS ALMOST BANKRUPT, ***which*** **is not surprising.**

7. We've GOT FRIDAY AFTERNOONS OFF, ***which*** **is very good.**

In the example 6 and 7, the sentential relative clause is only formed by the relative pronoun *which* rather than any other relative pronouns such as *that, who, where* or *when*. The relative pronoun *which* refers to the whole sentence. For instance, in the example 6, ***which*** **is not surprising** refers to the whole sentence and it means 'and this is not surprising'. In the example 7, ***which*** **is very good** refers to the whole sentence and it means 'and this is very good'.

2.7.2 Turkish Relative Clauses

Relative clauses are complex adjectival phrases modifying NPs. There are three participle suffixes, *-(y)An, -DIK, or –(y)AcAK*, used to construct relative clauses in Turkish. These suffixes correspond to the relative pronouns 'who', 'which', 'that', 'whom', 'whose', 'where', etc in English (Göksel & Kerslake, 2005, p. 381).

8. [burada sat-ıl-***an***] kitaplar (cf. Kitaplar burada satılıyor.)
   here sell-PASS-PART book – PL
   'the books [(***which*** are) sold here]'

The RC- forming three participles, *-(y)An, -DIK*, and *–(y)AcAK,* have different functions while relativizing NPs, such as relativization of different elements of a sentence such as subject , object, oblique object or indicating different tenses. In the table 3 we can see the summary of strategies of relativization:

Table 3

Summary of Strategies of Relativization in Turkish (adapted from Göksel & Kerslake, 2005, p. 387)

Relativizing subjects	- (y)An
Relativizing objects	-DIK / - (y) AcAK
Relativizing oblique objects	-DIK / -(y)AcAK
Relativizing adverbials	(y)An
Relativizing possessors:	
(a) which are part of subjects	(y)An
(b) which are not part of subjects	(y)An or -DIK / - (y)AcAK
Relativizing possessed constituents	
(a) which are part of subjects	(y)An
(b) which are not part of subjects	-DIK / -(y)AcAK

In the examples 9, 10, 11 and 12, there are sentences that exemplify the relativization in Turkish.

9. **Oyuncak-lar-ın-ı kır-*an*** küçük kız
   toy- PL-3SG- POSS. ACC break – PART little girl
   the little girl who has broken her toy
10. **her gün okul-da gör-*düğ*-üm** kız
    every day school- LOC see PART – 1SG- POSS girl
    the girl whom I see at school every day
11. **anne-si-y-le tanış- *acağ*- ım kız**
    mother – 3SG. POSS. COM meet- PART- 1SG.POSS girl

the girl whose mother I'm going to meet

12. Fatma'nın **yarın gör-*eceğ*-i** film
    Fatma- GEN. tomorrow see – PART- 3SG. POSS film
    the film that Fatma is going to see tomorrow

2.7.2.1. Restrictive and Non-Restrictive Relative Clauses in Turkish

Turkish has restrictive and nonrestrictive relative clauses. Restrictive relative clauses are those which restrict the denotation of the head noun. Non-restrictive relative clauses do not limit the denotation of the head noun, but they add extra, optional information about the head noun (Göksel & Kerslake, 2005, p. 388).

13. [Bugün yurda dön-**en *Cumhurbaşkanı*** saat 16'da bir basın toplantısı düzenleyecek.

'The president, [who is returning to the country today], will hold a press conference at 4 p.m.'

In the example 13, it is clear that in Turkish there is no formal distinction between restrictive and nonrestrictive relative clause such as punctuation as commas in English.

2.7.3. Japanese Relative Clauses

The general syntactic structure of Japanese relative clause comprises of the noun phrase (NP) head's being modified by a verb phrase (VP) clause body in comprising the overall noun phrase. Formally there is no distinction between N modification by V and adjective and relative clauses. In both cases, a V/adj/AN, or a clause ending in V/Adj./AN in normal word order is simply placed before a following N, i.e. unlike in English there are no relative pronoun or changes in word order. As there are not any relative pronoun or any other indication of the case relation between the modifying section and the N it modifies, it is sometimes only from the context or knowledge of the word that it can be determined (Kaiser, Ichikawa, Kobayashi & Hilofumi, 2001, p. 447).

As we can see in the examples 14, 15, 16 and 17, when the noun modification ends in **–ta**, it has different meaning depending on whether **–ta** indicates a state, or

completion of action. In the first example, **-ta** indicates completion, yet in the second example it indicates a state describing *what sort of burglars* they are (Kaiser et al. 2001, p. 154).

14. tökyö kara ki**ta** iguchi desu.
    I'm Iguchi, [who has come] from Tokyo.
15. saikin takuhaibin o yosoo**ta** götö ga deru sö da.
    Recently, there are supposed to be burglars [who are dressed as a home delivery service].
16. kowaku natta shufu wa hobo zengaku o shiharat**ta.**
    The housewife, [who got frightened], paid more or less the whole amount.
17. itsumo to kawaranai yuga**ta** data.
    [It] was an enemy [that was] no different from usual

Based on the similarity between Turkish and Japanese relative clauses, it can be said that there is no formal distinction between restrictive and non-restrictive relative clauses both in Turkish and Japanese. However, in English, non-restrictive relative clauses are set off by comma intonation, and written between commas.

CHAPTER 3

METHODOLOGY

3.1. Introduction

The methodology for this study is based on descriptive research design. This study focuses on examining the use of relative clauses with the relative pronoun *which* in Turkish university students' written English as to underuse, overuse and error types and also is to compare it with the use in Japanese university students' written English and in both native English and Turkish written texts use.

3.2. Research Design

This study is based on mixed type of descriptive quantitative research design and contrastive interlanguage analysis on the basis of data driven methodology to describe the underuse, overuse and errors of relative pronoun *which* in Turkish university students' argumentative essays in English. The significant differences between corpora in terms of underuse overuse and errors within the written texts are described and outlined.

3.3. Material

In the present study, a corpus based data has been utilized for the purpose of analysis. We have utilized two learner corpora: TICLE, the Turkish sub-corpus of International Corpus of Learner English and JPICLE, the Japanese sub-corpus of International Corpus of Learner English. Furthermore, we have utilized two native corpora: LOCNESS and TUC. The TICLE and JPICLE corpora present the learner English, namely nonnative speakers (NNS) of English. The TICLE is a collection of Turkish adult learners' argumentative essays in English. Similarly, the JPICLE is a collection of Japanese adult learners' argumentative essays in English. TICLE has been analyzed, for the specific aim of this study, to describe how the relative pronoun *which* in English the Turkish adult learners of English perform in their written performances. TICLE and JPICLE data have been compared with LOCNESS, the native form of the

same kind of corpus which includes American university students' argumentative essays. In addition, these two learner corpora have been compared with TUC, the native form of the same kind of corpus which includes Turkish speakers' argumentative essays. The descriptions of the corpora used in this study are presented below.

3.3.1. Corpora Used in the Study

In this study, we have set out to examine the frequency and use of the relative pronoun *which* in both restrictive and nonrestrictive relative clauses in learner and native corpora. Our main data sources are Turkish International Corpus of Learner English (TICLE) and Japanese International Corpus of Learner English (JPICLE), which are both the sub corpora of International Corpus of Learner English (ICLE) and two native corpora Louvain Corpus of Native English Essays (LOCNESS) and Turkish University Corpus (TUC). ICLE collection project has been directed by the Universite Catholique de Louvain in Belgium. The Turkish sub corpora TICLE, collected from the English language department students of Çukurova, Mersin and Hatay Mustafa Kemal universities, has been coordinated by the researchers from Çukurova University in Turkey. The Japanese sub corpora has been collected and coordinated by researchers from Showa Women's University in Japan. TICLE consists of 280 texts and 199, 532 words in total. JPICLE consists of 366 texts and 198, 241 words in total. As regards to the native corpora used in the study, LOCNESS consists of 175 texts and 149, 501 words. TUC consists of 183 texts and 107, 876 words. The total word numbers of TICLE and JPICLE have been balanced based on the word number of LOCNESS. The following table displays the size of corpora as regards to the selected texts and number of words.

Table 4

The Size of the Corpora

Corpora	**Selected Texts**	**Number of Words**
TICLE	208 – 280	149, 304
JPICLE	271 – 366	149, 368
LOCNESS	175 – 175	149, 501
TUC	183 – 183	107, 876

As can be seen from the table 4, the corpora used in the study are of considerably similar size and produced under similar conditions and are the product of the similar argumentative writing essays. The Turkish learner corpus (TICLE) was produced by advanced Turkish university students of English. In a similar way, The Japanese learner corpus (JPICLE) was produced by advanced Japanese university students of English. On the one hand, the native English speaker corpus utilized in the study was taken out from the LOCNESS database and involves in similar argumentative topics written by the native speaker of American university students. Turkish University Corpus (TUC) was produced by the native Turkish university students and it also involves in similar argumentative topics.

3.3.2. ICLE

During the formation of ICLE design criteria, corpus status and the highly heterogeneous nature of learner language were primarily taken into consideration. Strict control was done on task and learner variables, or else the data would be difficult to interpret. As a matter of fact, the collection of clearly expressed samples of learner language can assist researches to have clear statements about learner errors.

Table 5

ICLE task and Learner Variables (adapted from Granger et al*. 2009, p. 3)*

International Corpus of Learner English	
Task Variables	**Learner Variables**
-Medium	-Age
-Genre	-Gender
-Field	-Mother tongue
-Length	-Region
-Topic	-Other FLs
-Task setting	-Stay in English speaking country
	-Proficiency Level

All the variables were collected via a learner profile questionnaire, which all learners were requested to fill in, and subsequently included in the database. The main

goal of the ICLE project was to gather learner productions which include in task variables such as medium (writing), genre (academic essay), field (general English) and length (between 500 and 1000 words). The following sections and tables display the task and learner variables of ICLE (Granger et al. 2009, p. 3).

3.3.2.1. Genre

The majority of the ICLE texts are argumentative essays, which make possible for research on discourse, lexical and grammatical aspect. The proportion of argumentative essays ranged from 79% to 100%.

Table 6

Proportion of Argumentative Essays (adapted from Granger et al*. 2009, p. 4)*

National Subcorpus	**Argumentative**
Bulgarian	100%
Chinese	100%
Czech	81%
Dutch	96%
Finnish	92%
French	85%
German	97%
Italian	34%
Japanese	100%
Norwegian	98%
Polish	99%
Russian	100%
Spanish	79%
Swedish	85%
Turkish	100%
Tswana	100%
***ICLE*v2**	**91%**

The proportions of argumentative essays in Turkish and Japanese subcorpora are both 100%. The nature of argumentative essays is discourse oriented and enables the researchers to explore grammar enquiries and lexical (Granger et al. 2009, p. 4).

Table 7

Average Essay Length (adapted from Granger et al. *2009, p. 6)*

National Subcorpus	**Average Length**
Bulgarian	663 words
Chinese	500 words
Czech	830 words
Dutch	893 words
Finnish	704 words
French	654 words
German	526 words
Italian	572 words
Japanese	542 words
Norwegian	668 words
Polish	641 words
Russian	832 words
Spanish	789 words
Swedish	564 words
Turkish	713 words
Tswana	384 words
***ICLE*v2**	**617 words**

All the essays are unabridged and have an average length of 617 words. However, some of the subcorpora such as Tswana by 384 words differentiate each other in the corpus. Turkish subcorpora have an average length of 713 words while Japanese subcorpora have an average length of 542 words (Granger et al. 2009, p. 6).

3.3.2.2. Topic

The ICLE essays contain a wide range of topics that can be exemplified as poverty, prison system, money, smoking ban, university education, feminism and politics. Table 8 presents the list of the ten most popular topics.

Table 8

Top Ten Essay Topics (adapted from Granger et al. *2009, p. 7)*

Essay Topic	**Number of Essays**
1. Some people say that in our modern world, dominated by science, technology and industrialization, there is no longer a place for dreaming and imagination. What is your opinion?	491 / 29% Bulgarian
2. Most university degrees are theoretical and do not prepare students for the real world. They are therefore of very little value.	249 / 22% Turkish
3. Poverty is the cause of the HIV/AIDS epidemic in Africa.	243 / 100% Tswana
4. Marx once said that religion was the opium of the masses. If he was alive at the end of 20th century, he would replace religion with television.	237 / 19% Russian
5. The prison system is outdated. No civilized country should punish its criminals: it should rehabilitate them.	176 / 32% Tswana
6. Discuss the advantages and disadvantages of banning smoking in restaurants.	156 / 100% Chinese
7. Discuss the advantages and disadvantages of using credit cards.	149 / 100% Chinese
8. Feminists have done more harm to the cause of women than good.	139 / 23% Russian
9. In the words of the old song ‘‘Money is the root of all evil’’.	133 / 22% Russian
10. In his novel ‘‘Animal Farm’’, George Orwell wrote ‘‘All men are equal; but some are more than others’’. How true is this today?	127 / 39% Bulgarian

As regards to the choice of topics on average, 22% of Turkish subcorpora are comprised of the topic university degree (Granger et al. 2009, p. 7).

3.2.2.3. Task Setting

This variable is related to three task conditions: timed or untimed, under exam conditions or not, and allowance for reference tools. Most of the ICLE essays are of untimed essay and not written under exam conditions. Nearly half of the essays by 48% have been written with the support of reference tools (Granger et al. 2009, p. 7).

3.3.2.4. Learner Variables

Granger et al. (2009, p. 7) state that age, gender, mother tongue background, region, knowledge of other foreign languages and time are regarded as six distinct variables. However, two variables learner context and proficiency level are comparatively unclear.

3.3.2.4.1. Age

All the essays are comprised of learners who are university undergraduates in their twenties. As shown in Table 9, the average age of Japanese subcorpus is 20.06 and the average age of Turkish subcorpus is 22.08. The whole average age of ICLE is 22.30.

Table 9

Learners' Age in JPICLE and TICLE (adapted from Granger et al. *2009, p. 8)*

National Subcorpus	**Average Age**
Japanese	20.06
Turkish	22.08
***ICLE*v2**	**22.30**

3.3.2.4.2. Gender

As regards to the proportion of female and male students, it has not been possible to form a well-balanced corpus. As we can see in Table 10, 76% of the ICLE data have been formed by female learners.

Table 10

Learners' gender (adapted from Granger et al. *2009, p. 9)*

National Subcorpus	**Percentage Female**
Bulgarian	83%
Chinese	64%
Czech	72%
Dutch	73%
Finnish	85%
French	88%
German	78%
Italian	92%
Japanese	73%
Norwegian	74%
Polish	80%
Russian	84%
Spanish	86%
Swedish	77%
Turkish	81%
Tswana	60%
***ICLE*v2**	**76%**

The percentage female of Japanese subcorpus is 73% while the percentage female of Turkish subcorpus is 81%. These figures present the noticeable percentage of female (Granger et al. 2009, p. 9).

3.3.2.4.3. Mother Tongue

In the corpus there are 16 different mother tongue backgrounds: Bulgarian, Chinese, Czech, Dutch, Finnish, French, German, Italian, Japanese, Norwegian, Polish, Russian, Spanish, Swedish, Turkish and Tswana. The language background of learners has also been stored which other languages are spoken at home. These languages have been listed in the usage frequency of first, second and third language spoken at home (Granger et al. 2009, p. 9).

3.3.2.4.4. Region

Table 11

Languages Spoken in More Than One Country

Language	**The places where it is spoken**
Chinese	Mainland China and Hong Kong
Dutch	Belgium and the Netherlands
German	Germany, Austria and Switzerland
Swedish	Sweden and Finland

The origin of learners' country is covered in this variable. The languages spoken in more than one country are available in the corpus. The researchers can utilize the information of region in then corpus as regards to the places where a language is spoken (Granger et al. 2009, p. 9).

3.3.2.4.5. Knowledge of other Foreign Languages

The learners' L2 could be affected by their mother tongue or by their knowledge of other foreign languages. The first foreign language in the ICLE is German by 32%, followed by French 27% (Granger et al. 2009, p. 10).

3.3.2.4.6. Time Spent in an English-Speaking Country

A great number of the learners by 45% have not been in an English speaking country, 23% have been in an English speaking country for 3 months or more and 19% have been for less than 3 moths. 13% is not known (Granger et al. 2009, p. 10).

3.3.2.4.7. Learning Context

All the learners in the corpus data have learned English in a non-English-speaking country. Thus, English is a foreign language rather than a second language. Initially we have mentioned that learning context and proficiency are comparatively unclear and fuzzy. The learners' construction of English may differentiate between the countries in the corpus. For instance, in Norway English language is becoming a second language rather than a foreign language. Although this variable seems unclear and fuzzy, it is clear that all the learners represented in the ICLE corpus have learnt English primarily in a classroom setting (Granger et al. 2009, p. 11).

3.3.2.4.8. Proficiency

The learners are university undergraduates in English in their third or fourth year. The corpus varies in terms of proficiency level, which ranges from higher intermediate to advanced. A random sample of 20 essays from each of the 16 subcorpora was sent to a professional rater who determined the levels on the basis of the common European Framework of Reference for Languages (CEF) descriptors for writing. Table 12 presents CEF results including 20 essays per subcorpus.

Table 12

CEF Results – 20 Essays Per Subcorpus (adapted from Granger et al. *2009, p. 9)*

Mother tongue	**B2 (and lower)**	**C1**	**C2**	**Total**
Bulgarian	2	16	2	20
Chinese	19	1	0	20
Czech	11	9	0	20
Dutch	1	11	8	20
Finnish	3	8	9	20
French	3	11	6	20
German	1	12	7	20
Italian	10	9	1	20
Japanese	18	2	0	20
Norwegian	8	7	5	20
Polish	1	12	7	20
Russian	3	15	2	20
Spanish	12	8	0	20
Swedish	0	14	6	20
Tswana	18	0	2	20
Turkish	16	4	0	20
Total	**126**	**139**	**55**	**320**

While 60% of the sample essays were rated as advanced C1 or C2, the proportion of advanced is higher in some subcorpora such as Swedish, Dutch, Polish, German and French. In Japanese subcorpus 18 essays are B2 and lower, 2 essays are C1 and there are not any C2 essays. Similarly, in Turkish subcorpus 16 essays are B2 and lower, 4 essays are C1 and there are not ant C2 essays. In total, 126 essays are B2 and lower, 139 essays are C1 and 55 essays are C2 (Granger et al. 2009, p. 12).

3.3.3. LOCNESS

In this study LOCNESS has been chosen as the reference corpus because there is a chance of comparability between ICLE and LOCNESS. Furthermore, LOCNESS has been used most commonly for comparison. For instance, Virtanen (1988) studied direct questions in argumentative student writing by comparing several NNS subcorpora of ICLE and LOCNESS. Altenberg and Tapper (1998) also studied the use of adverbial connectors in advanced Swedish learners' written English by comparing native corpus LOCNESS and NNS subcorpora of ICLE. LOCNESS was constructed by the Centre for English Corpus Linguistics at Catholic University of Louvain, Belgium and made available for public use in 1998. The texts of the corpus include argumentative essays. The length of essays is around 500 words. The age of students is mostly between 17 and 23. The texts cover a very wide range of topics from social problems, water pollution, nuclear power, violence, sex, gender issues. In this study the native corpus called LOCNESS consists of 175 texts and 149, 501 words.

3.3.4. TUC

TUC which is the native form of the same kind of corpus includes Turkish speakers' argumentative essays. TUC includes argumentative essays. The length of essays is around 550 words. The age of students is mostly between 17 and 24. The texts cover a very wide range of topics from social problems, water pollution, nuclear power, violence, sex, gender issues. It consists of 183 texts and 107, 876 words.

3.3.5. Software

In this study the corpora mentioned have been investigated by *ICLE* software and *WORDSMITH tools Scott version.* Turkish and Japanese interlanguage learner data have been investigated through their own database program *ICLE* itself. However, the native corpora LOCNESS and TUC have been investigated using *WORDSMITH tools* software. Both *ICLE* and *WORDSMITH* software have been designed for program for text analysis that retrieve word lists and concordances from the written texts coded in *txt* format by frequency and by alphabetical order. *WORDSMITH* provides quite practical outlook for concordances, wordlists and keywords.

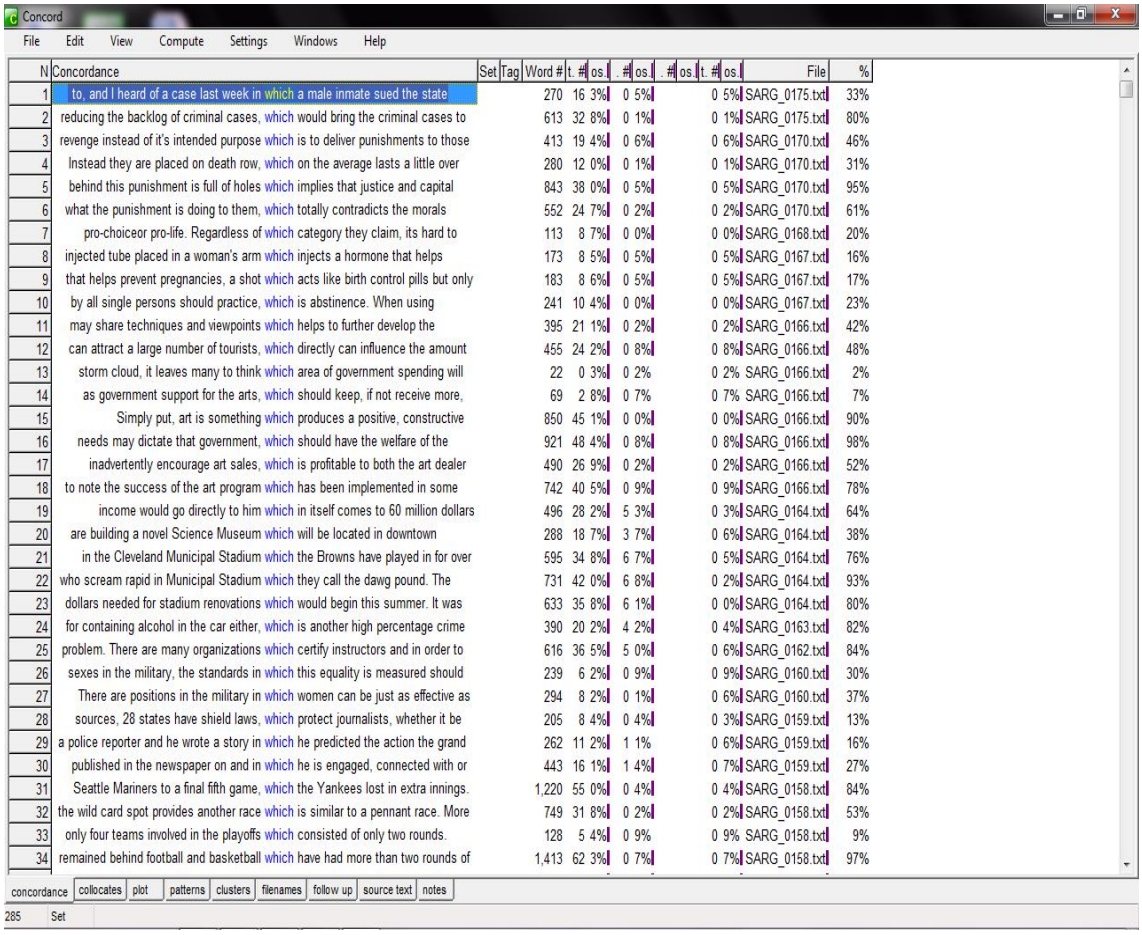


Figure 1. A screenshot of the relative pronoun *which* (from LOCNESS) by WordSmith

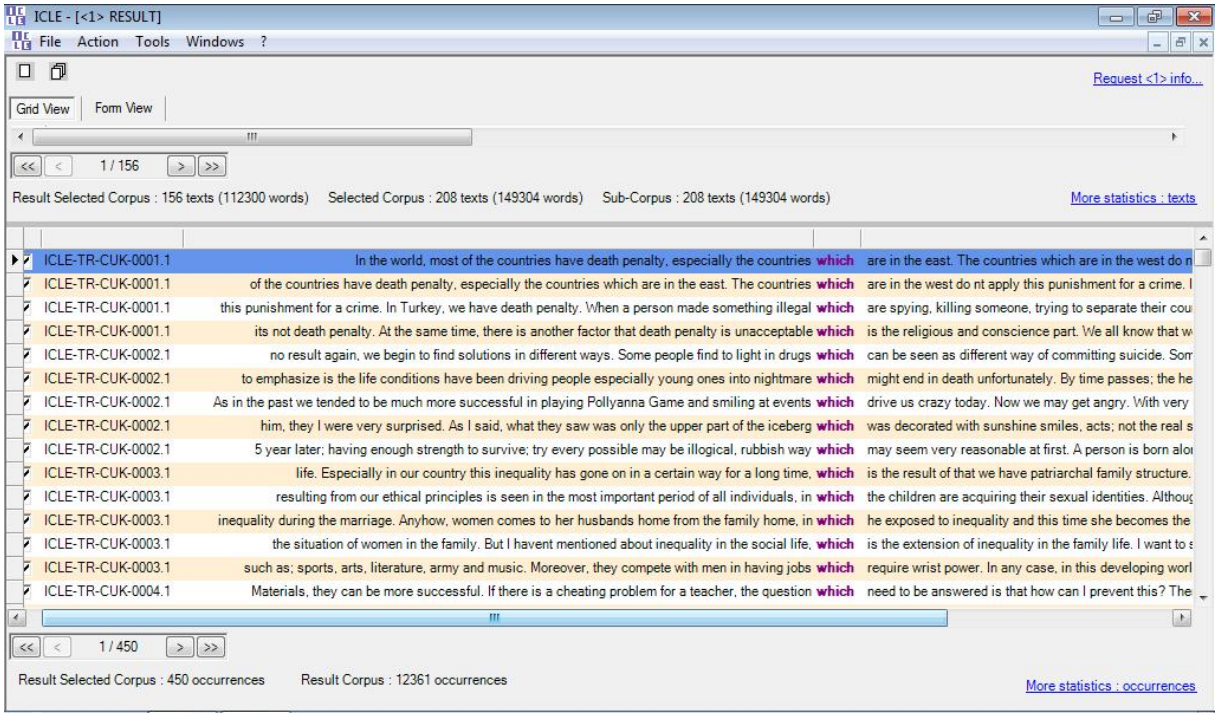


Figure 2. A screenshot of the relative pronoun *which* (from TICLE) by ICLE

These software packages are able to identify key words to retrieve the sentences related to the data under investigation.

For instance, the sentences that include in the relative clauses with the relative pronoun *which* from TICLE and JPICLE in order are shown as it is in the examples below:

18. Most of the countries have death penalty, especially in the countr**ies** *which* **are** in the east….. (ICLE-TR-CUK-0001.1)
19. There are numbers of school**s** *which* **provide** bilingual education, and all schools are required by law…… (ICLE-JP-KO-0002.2)

After having the concordance lines both from Turkish and Japanese learner English corpora, we have analyzed each relative pronoun *which* sentence based on categories. For instance, we have firstly grouped the sentences as restrictive and non-restrictive relative clauses. Secondly, we have analyzed these sentences according to the errors learners have done such as concord problem, sentential errors in non-restrictive relative clauses, relative pronoun choice errors.

3.4. Data Analysis

In this study, we aim to search the use of relative clauses with the relative pronoun *which* in Turkish university students' written English as to underuse, overuse and error types and compare it with the use in American university students' argumentative essays. The statistical analyses carried out in the study have been done with the use of statistical software called Statistical Package of Social Sciences (SPSS for Windows*, version 11.0*) under the guidance of a statistician. Afterwards, the SPSS values of relative pronoun which in all corpora in the study were compared among four of them with Chi-square analysis to find out whether there is a statistical significance.

CHAPTER 4

RESULTS AND DISCUSSION

4.1. Introduction

In this section, data obtained from both Turkish and Japanese learner corpora and native speaker American university students and Turkish university students' corpora examining the use of relative clauses with the relative pronoun *which* as to under and overuse and errors are analyzed and discussed.

4.2. Analysis of the Data from Corpora

The overall use of the relative pronoun *which* has been presented as regards to frequency and percentages based on the total tokens. In addition to include the errors in the analysis, the correct and incorrect usages of learner corpora data have also categorized. The four corpora in the study have been calculated by means of *SPSS*. Afterwards, the values between two groups, namely two corpora, were compared via Chi-square analysis to examine significance and underuse/overuse.

Table 13

The Overall Use of the Relative Clauses with the Relative Pronoun Which in TICLE, JPICLE, LOCNESS and TUC.

Groups	***f***	***%***	**Total Tokens**
TICLE	362	0,24	149,304
JPICLE	311	0,20	149,368
LOCNESS	200	0,13	149,501
TUC	1597	1,48	107,876

As can be seen in Table 13, in TICLE there are 362 usages referring to 0, 24 percent in the total token. In JPICLE there are 311 usages comprising of 0, 20 percent of the whole token in the subcorpus. Furthermore, the native speaker Turkish university

students' and native speaker American university students' written argumentative essays in their own mother tongue have been analyzed in this study. When we look at the results obtained from these native corpora, in LOCNESS there are 200 usages, which equal to 0, 13 percent of the total token and in TUC there are 1597 usages, which equal to 1, 48 percent of the total token.

The table 13 also presents that the Turkish learners use more the relative pronoun *which* in their essays than the native speaker American students (362 vs. 200 usages). This indicates that Turkish university students tend to overuse the relative pronoun *which* in their written English. Interestingly, Turkish university students use fewer the relative pronoun *which* than Turkish students do when writing in their native language. To seek answers for our research questions, Chi-square tests were performed for each variable. In this respect, we attempted to seek answers to two of our research questions which are Do Turkish university students use the relative pronoun *which* to the same extent as native speaker American university students ? And is there a significant difference between underuse and overuse of the relative pronoun *which* and is this due to mother tongue influence or an interlanguage universal? Table 14 illustrates the results obtained from Chi-square tests of comparable corpora types.

Table 14

The Comparison among Turkish and Japanese Learner Corpora and the Native Corpora in the Relative Clauses with the Relative Pronoun Which

	f	**(p <.05)**
Groups		
TICLE - JPICLE	362 -311	,001*
TICLE - LOCNESS	362- 200	,000*
TICLE - TUC	362- 1597	,000*
JPICLE - LOCNESS	311- 200	,417
LOCNESS - TUC	200 - 1597	,000*

a 0 cells (, 0%) have expected count less than 5.

b Computed only for a 2×2 table

In the light of the numerical data provided in Table 14, the result between the TICLE and LOCNESS reveals a significant difference (362 vs. 200). In a similar way,

there is a significant difference between TICLE and TUC (362 vs. 1597). Moreover, the significant difference has been found between TICLE and JPICLE (362 vs. 311). While the result between JPICLE and LOCNESS does not reveal a significant difference, the result between JPICLE and TUC shows a significant difference. There is also a significant difference between LOCNESS and TUC (200 vs. 1597). To sum up, Turkish university students' overall use in the relative pronoun *which* presents difference from the use of American university students. This answers our research question whether Turkish university students use the relative pronoun *which* to the same extent as native speaker American university students. It can be seen that Turkish learners tend to overuse the structure relative pronoun *which* than JPICLE and LOCNESS (362 vs. 311), (362 vs. 200). This also suggests that the reason of overuse in Turkish university students might be due to the influence from mother tongue Turkish.

Table 15

The Overall Use of the Restrictive Relative Clauses with the Relative Pronoun Which in TICLE, JPICLE, LOCNESS and TUC.

Groups	***f***	***%***	**Total**
TICLE	290	80, 11	362
JPICLE	220	70, 73	311
LOCNESS	126	63	200
TUC	1597	100	1597

The table 15 shows the occurrences of the relative pronoun *which* in restrictive relative clauses. In TICLE there are 290 usages out of 362 tokens, which is equal to 80, 11 percent. In JPICLE there are 220 usages out of 311 tokens, which is equal to 70, 73 percent. In LOCNESS there are 126 usages out of 200 tokens, which is equal to 63 percent. In TUC there are 1597 usages out of 1597. Therefore, it can be said that Turkish university students tend to overuse the relative pronoun *which* in restrictive relative clauses (290 vs. 126 usages). On the other hand, in TICLE, Turkish university students overused the relative pronoun *which* than the students of other learner corpora JPICLE (290 vs. 220 usages). When compared to the other learner corpora JPICLE and LOCNESS, it can be seen that there are only 220 usages out of 311 tokens in the

relative pronoun *which* in restrictive clauses, which does not reveal a significant difference as it does between TICLE and LOCNESS. This suggests that Turkish university students' overuse of restrictive relative clauses with the relative pronoun *which* might be due to the tendency of overusing relative pronoun *which* in mother tongue Turkish. The following table 16 represents the results of the Chi-square tests of comparable corpora in the relative pronoun *which* in restrictive relative clauses.

Table 16

The Comparison among Turkish and Japanese Learner Corpora and the Native Corpora in the Restrictive Relative Clauses with the Relative Pronoun Which

	f	**(p <.05)**
Groups		
TICLE- JPICLE	290 - 220	,000*
TICLE - LOCNESS	290 - 126	,000*
TICLE- TUC	290 - 1597	,000*
JPICLE - LOCNESS	220 - 126	,175
LOCNESS - TUC	126 - 1597	,000*

a 0 cells (, 0%) have expected count less than 5.

b Computed only for a 2×2 table

On the basis of the outcomes the test has yielded, TICLE and JPICLE reveal a significant difference by the occurrences 290 vs. 220. Furthermore, there are significant differences among TICLE, LOCNESS and TUC. For instance, there is a significant difference between TICLE and LOCNESS (290 vs. 126). There is also a significant difference between TICLE and TUC (126 vs. 1597). However, no significant difference has been found between JPICLE and LOCNESS. The test results suggest that the Turkish learners might be influenced from their mother tongue Turkish.

Table 17

The Overall Use of Non- Restrictive Relative Clauses with the Relative Pronoun Which in TICLE, JPICLE, LOCNESS and TUC.

Groups	***f***	***%***	**Total**
TICLE	72	19, 89	362
JPICLE	91	29, 26	311
LOCNESS	74	37	200
TUC	0	0	1597

As the table 17 illustrates, in TICLE, there are 72 usages out of 362 tokens, which is equal to 19, 89 percent. In JPICLE there are 91 usages out of 311, which is equal to 29, 26 percent. In LOCNESS there are 74 usages out of 200, which is equal to 37 percent. These results suggest that Turkish learners underused the non-restrictive relative clauses with the relative pronoun *which* (72 vs. 74). However, it can be said that there is a similarity between JPICLE and the native speaker American university students' corpus LOCNESS. When comparing to the overall occurrences of restrictive and non restrictive relative clauses with the relative pronoun *which*, we come to conclusion that there are more restrictive clauses than non-restrictive clauses in four types of corpora. As there is no formal distinguish between restrictive and non-restrictive relative clauses, in TUC there are not any usages of non-restrictive relative clauses with the relative pronoun *which*.

Table 18

The Comparison among Turkish and Japanese Learner Corpora and the Native Corpora in the Non-Restrictive Relative Clauses with the Relative Pronoun Which

	f	**(p <.05)**
Groups		
TICLE - JPICLE	72 - 91	,247
TICLE - LOCNESS	72 - 74	,882
TICLE - TUC	72 - 0	,000*
JPICLE - LOCNESS	91 - 74	,207
LOCNESS - TUC	74 - 0	,000*

a 0 cells (, 0%) have expected count less than 5.

b Computed only for a 2×2 table

In the light of numerical data provided in the table 18, the result between TICLE and TUC reveals a significant difference (72 vs. 0). Similarly, there is also a significant difference between two native corpora LOCNESS and TUC (74 vs. 0). In non-restrictive relative clauses, Japanese learners tend to when compared with American students (91 vs. 74) and Japanese learners also tend to overuse when compared with Turkish learners (91 vs. 72). Nevertheless, the results between TICLE and JPICLE, TICLE and LOCNESS, and also between JPICLE and LOCNESS do not reveal any significant differences.

Table 19

The Overall Errors in the Relative Clauses with the Relative Pronoun Which in TICLE and JPICLE

Groups	***f error***	***% error***	**Total**
TICLE	85	23, 48	362
JPICLE	133	42, 76	311

In table 19, there are 85 errors out of 362 tokens, which equals to 23, 48 percent of the total token. However, there are 133 errors out of 311 tokens, which equals to 42, 76 percent of the total token. The occurrences of overall errors between TICLE and

JPICLE reveal that there are more errors in JPICLE than TICLE (133 vs. 85). The results suggest that Japanese learners tend to make more errors than Turkish learners (23, 48 percent vs. 42, 76).

Table 20. *The Overall Errors in the Restrictive Relative Clauses with the Relative Pronoun Which in TICLE and JPICLE*

Groups	***f error***	***% error***	**Total**
TICLE	25	8, 62	290
JPICLE	58	26, 36	220

As the table 20 illustrates, the occurrences of errors between two learner corpora TICLE and JPICLE are 25 vs. 58. That is to say, while there are 25 errors out of 290 by 8, 62 percent of the total tokens in TICLE, there are 58 errors out of 220 by 26, 36 percent of the total tokens. Similar to the overall errors of the relative clauses with the relative pronoun *which*, Japanese learners tend to make more errors than Turkish learners in the restrictive relative clauses with the relative pronoun *which*.

Table 21

The Comparison of Errors among Turkish and Japanese Learner Corpora and the Native Corpora in the Restrictive Relative Clauses with the Relative Pronoun Which

	f error	**(p <.05)**
Groups		
TICLE - JPICLE	25 - 58	,395
TICLE - LOCNESS	25	,000*
TICLE - TUC	25	,000*
JPICLE - LOCNESS	58	,000*

a 0 cells (, 0%) have expected count less than 5.

b Computed only for a 2×2 table

According to the results obtained from the tests, it can be seen that there is a significant difference between TICLE and LOCNESS. Furthermore, the result between

TICLE and TUC reveals a significant difference. There is also a significant difference between the other learner corpus JPICLE and LOCNESS. On the other hand, there is no significant difference between two learner corpora TICLE and JPICLE. These figures present that the errors made by Turkish learners might be affected by their mother tongue since there is not a significant difference between two learner corpora.

Table 22

The Overall Errors in the Non-Restrictive Relative Clauses with the Relative Pronoun Which in TICLE and JPICLE

Groups	***f error***	***% error***	**Total**
TICLE	60	83, 33	72
JPICLE	75	16, 66	91

As can be seen in the table 22, in TICLE while there are 60 errors out of 72 tokens by 83, 33 percent, in JPICLE there are 75 errors out of 91 tokens by 16, 66 percent. Similar to the overall errors of the restrictive relative clauses with the relative pronoun *which*, Japanese learners tend to make more errors than Turkish learners in the restrictive relative clauses with the relative pronoun *which* (75 vs. 60). However, the percentages of the total tokens in TICLE and JPICLE vary into each other (83, 33 percent vs. 16, 66 percent).

Table 23

The Comparison of Errors among Turkish and Japanese Learner Corpora and the Native Corpora in Non-Restrictive Relative Clauses with the Relative Pronoun Which

	f error	**(p <.05)**
Groups		
TICLE - JPICLE	60 - 75	,391
TICLE - LOCNESS	60	,000*
TICLE - TUC	60	,000*
JPICLE - LOCNESS	75	,000*

a 0 cells (, 0%) have expected count less than 5.

b Computed only for a 2×2 table

As the test results in table 23 illustrate, there is a significant difference between TICLE and LOCNESS. Moreover, the result between TICLE and TUC reveals a significant difference. Similarly, there is a significant difference between JPICLE and LOCNESS. However, the result between TICLE and JPICLE does not reveal any significant difference (60 vs. 75).

4.3. Analyses of Errors in TICLE Learner Corpus

In this section, the samples of categorized Turkish university students' errors have been presented and outlined. In our study, four types of error such as restrictive, non-restrictive, sentential and choice of relativizer occurred. Biber (2006, p. 76) claims that most learners have great difficulty in producing such structures as relative clauses with relative pronoun which due to its embodying linguistic complexity as restrictive and non-restrictive. The errors of Turkish university students are illustrated in the following sections.

4.3.1. Errors in Relative Pronoun *Which* in Restrictive Relative Clauses

The errors in the restrictive relative clauses with the relative pronoun *which* have been presented in this section. The main function of relative clauses in English is to modify a noun phrase. Leech and Svartvik (2002) state that the linking is achieved with a back-pointing element generally a relative pronoun. The relative pronouns in English point back to the head of the noun phrase which is called the antecedent (p. 383). In this study, the focus is the relative pronoun *which*. Although Turkish students seemed to be successful in pointing back to the right antecedent, they tended to make errors of antecedent and relative clause concord.

4.3.1.1. Concord Errors in the Relative Pronoun *which* in Restrictive Relative Clauses

Göksel and Kerslake (2005, p. 380) express that relative clauses are complex adjectival constructions that modify noun phrases and the most common type of relative clause is non-finite and contains participle suffixes such as *– (y) An*, *-DIK*, or *– (y) AcAK* corresponding to the relative pronouns 'who', 'which', 'that' , 'whom', 'whose' ,

'where', etc in English. As Turkish has participle suffixes, the noun of the noun phrase that a relative clause modifies namely the head noun can have several different relationships with the relative clause. Thus, Turkish learners might tend to make confuse the relationship between the head noun and the relative clauses. This also might be the explanation of mother tongue influence.

20. The other fact of suicide is financial problem**s** *which* especially **treats** the public of developing countries (ICLE-TR-CUK-0037.1).

In the example 20, the relative clause is '*which especially **treats** the public of developing countries'* and *the* relative pronoun '*which'* points back to the antecedent '*financial problem**s'***. However, this antecedent '*financial problem**s'*** is plural one and the relative clause should follow it with a relative clause like '*which especially **treat** the public of developing countries'*.

21. S/he can go best school**s** *which* **has** much opportunity, and qualified education (ICLE-TR-CUK-0056.1).

In the example 21, the relative clause is '*which has much opportunity'* and *the* relative pronoun '*which'* points back to the antecedent '*school**s'***. However, this antecedent '*school**s'*** is plural one and the relative clause should follow it with a relative clause like '*which have much opportunity, and qualified education'*.

22. Tutors only full the student's mind with mostly unnecessary thing**s** *which* only **works** with the student job (ICLE-TR-CUK-0064.1).

In the example 22, the relative clause is '*which only works with'* and *the* relative pronoun '*which'* points back to the antecedent '*unnecessary things'*. However, this antecedent '*unnecessary things'* is plural one and the relative clause should follow it with a relative clause like '*which only work with the student job'*.

23. Political and historical deficien**cies** *which* **is** responsible for universities to be inefficient in preparing the students for the real world (ICLE-TR-CUK-0068.1).

In the example 23, the relative clause is '*which is responsible for universities'* and *the* relative pronoun '*which'* points back to the antecedent '*Political and historical deficiencies'*. However, this antecedent '*Political and historical deficiencies'* is plural one and the relative clause should follow it with a relative clause like '*which are responsible for universities to be inefficient'*.

24. Today, it's not different as well political trick**s** *which* **causes** people's lack of attention on politics and prevent them to see the realities (ICLE-TR-CUK-0105.1).

In the example 24, the relative clause is '*which causes people's lack of attention'* and *the* relative pronoun '*which'* points back to the antecedent '*Political tricks'*. However, this antecedent '*Political tricks'* is plural one and the relative clause should follow it with a relative clause like '*which cause people's lack of attention'*.

25. Girls can easily read the book**s** *which* **is** peculiar to boys (ICLE-TR-CUK-0113.1).

In the example 25, the relative clause is '*which is peculiar to boys'* and *the* relative pronoun '*which'* points back to the antecedent '*the books'*. However, this antecedent '*the books'* is plural one and the relative clause should follow it with a relative clause like '*which are peculiar to boys'*.

26. There was a boy suffering from **a disease** *which* **are not** possible to be cured (ICLE-TR-CUK-0134.1).

In the example 26, the relative clause is '*which are not possible to be cured'* and *the* relative pronoun '*which'* points back to the antecedent '*a disease'*. However, this antecedent '*a disease'* is singular one and the relative clause should follow it with a relative clause like '*which is not possible to be cured'*.

27. Most university degree**s** *which* **is** a course of study at university or college are theoretical and do not prepare students (ICLE-TR-CUK-0172.1).

In the example 27, the relative clause is '*which is a course of study at university or college'* and *the* relative pronoun '*which'* points back to the antecedent '*most university degrees'*. However, this antecedent '*most university degrees'* is plural one and the relative clause should follow it with a relative clause like '*which are a course of study at university or college'*.

28. The other word, most universities give so much **detail** *which* **are not** important, and have no value in life (ICLE-TR-MER-0031.3).

In the example 28, the relative clause is '*which are not important'* and *the* relative pronoun '*which'* points back to the antecedent '*detail*'. However, this antecedent '*detail'* is singular one and the relative clause should follow it with a relative clause like '*which is not important, and has no value in life'*.

4.3.2. Errors in relative pronoun *which* in Non-restrictive Relative Clauses

Non-restrictive relative clauses do not limit the denotation of the head noun, but they add extra, optional information about the head noun. In Turkish there is no formal distinction between restrictive and nonrestrictive relative clause such as punctuation as commas in English or by the choice of relativizer (Göksel & Kerslake, 2005, p. 378). On the other hand, in English especially in written texts non-restrictive relative clauses are shown separately in written language by punctuation (Greenbaum & Quirk, 2003, p. 370). Accordingly, Turkish university students tended to make errors in non-restrictive relative clauses by adding punctuation without determining whether the relative clauses with pronoun *which* has the function of adding extra or optional information about the head noun.

4.3.2.1. Punctuation errors in the relative pronoun which in Non-restrictive relative clauses

29. If men use their disadvantage, *which* is due to creation of them for their bad desires (ICLE-TR-CUK-0005.1).

As can be seen in the example 29, the non-restrictive relative clause is created by the learner such as '*which is due to creation of them for their bad desires'* However, the relative clause in this example should not add extra or optional information about the antecedent '*disadvantage*'. This antecedent needs to be defined and the relative clause in this example does not need to be written by punctuation.

30. They need to feel the love of opposite sex, *which* is different from the family love (ICLE-TR-CUK-0036.1).

In the example 30, the relative clause should not add extra or optional information about the antecedent '*the love of opposite sex*'. It should not be separated by punctuation.

31. For example, the owner of a factory, *which* gives out industrial wastes, wouldn't do anything to prevent these wastes (ICLE-TR-CUK-0052.1).

In the example 31, the antecedent '*a factory*' should be explained. The relative clause in this example should have defined the head noun.

32. In my opinion the knowledge, *which* is learnt theoretically, is exposed to be forgotten and its permanence is very little value (ICLE-TR-CUK-0065.1).

In the example 32, the relative clause in this example does not add extra or optional information about the antecedent '*the knowledge*'. This antecedent should be defined and the relative clause in this example should not be written by punctuation.

33. If people see a deer, *which* had been hit by a car, and is in terrible pain, they will kill it (ICLE-TR-CUK-0080.1).

In the example 33, the antecedent '*a deer*' needs to be defined since it is not possible to determine which deer it is. Thus, the relative clause in this example should have been constructed as restrictive relative clause.

34. A baby, *which* we think that brings happiness, may bring unhappiness (ICLE-TR-CUK-0098.1).

Similar to the example 34, the antecedent '*a baby'* needs to be defined. Therefore, the restrictive relative clause should have been preferred.

35. The studies, *which* to a great extent are theoretical, are not totally invaluable as some people suggest (ICLE-TR-MER-0022.1).

In this example, the antecedent '*the studies'* are too general without defining with a restrictive relative clause. Thus, there is a necessity to limit the denotation of antecedent.

Unlike the examples presented previously, in the example 36, Turkish university students should have added the punctuation to separate the non-restrictive relative clause.

36. It is stated very clearly in *Hippocrates oath which* is taken by all doctors while beginning to their job (ICLE-TR-CUK-0140.1).

The antecedent in this example '*Hippocrates oath*' is a proper noun and it is already known that it is taken by doctors. Thus, the student should have regarded this relative clause as a non-restrictive relative clause.

4.3.2.2. Errors in Sentential Relative Clauses

It is a special type of non-restrictive relative clause. It does not point back to a noun phrase but to a whole clause or sentence. Among relative pronouns, only the relative pronoun *which* is used in sentential relative clauses (Leech & Svartvik, 2002, p. 387). As we discussed beforehand, there is no formal distinction between restrictive and non-restrictive relative clauses in Turkish university students' mother tongue. Thus, as we can also see in the examples 37, 38, 39 and 40, there is a tendency to show punctuation confusion in sentential relative clauses.

37. You have no chance of getting a food or a dress or anything you want *which* mean a misery for a human (ICLE-TR-CUK-0048.1).

38. At first, men have physically privilege *which* means that they are more powerful than women (ICLE-TR-CUK-0102.1).

39. They themselves took some innocent person's life *which* caused them to be put there (ICLE-TR-CUK-0103.1).

40. You apply your ability on making research to all parts in your life *which* makes you a real researcher (ICLE-TR-MER-0026.3).

4.3.3. Choice of Relativizer Pronoun

In the example sentences of this section, Turkish students tend to combine noun phrase for people with the relative pronoun *which.* Although the antecedents are noun phrases for people, the learners tend to combine the relative clause with them. For instance, in the example 41, the antecedent is '*matchmaker*'. However, the learner tends to combine it with the relative pronoun *which*. As Turkish language has participle suffixes in relative clauses, Turkish students might tend to confuse the relativizer pronoun and the head noun.

41. For instance, it can be your **matchmaker** *which* can find a person who fits hero qualifications or you can easily make your own friends (ICLE-TR-CUK-0028.1).

42. It is based on the people who help **the patient** *which* can be described as murderer or saver (ICLE-TR-CUK-0076.1).

43. They shouldn't marry **persons** *which* they don't love just for making their families happy (ICLE-TR-CUK-0086.1).

44. A student *which* makes research and improve himself from different aspects is not seen (ICLE-TR-MER-0002.1).

4.4. Analyses of Errors in JPICLE Learner Corpus

In this section, the samples of categorized Japanese university students' errors have been presented and outlined. Similar to the Turkish university students' corpus data, the errors occurred in the Japanese university students' data have been categorized

as restrictive, non-restrictive, sentential and choice of relativizer. The errors of Japanese university students are illustrated in the following sections.

4.4.1. Errors of Relative Pronoun *Which* in Restrictive Relative Clauses

Similar to the Turkish university students' errors in restrictive relative clauses, Japanese university students tended to confuse the noun phrase and relative clause. As we have discussed beforehand in the section 2.7.3., unlike in English there are no relative pronoun or changes in word order. As there are not any relative pronoun or any other indication of the case relation between the modifying section and the N it modifies, it is sometimes only from the context or knowledge of the word that it can be determined.

4.4.1.1. Concord errors in the Relative Pronoun *Which* in Restrictive Relative Clauses

45. In Tokyo, you can visit **several different places** *which* **is** very interesting in a day (ICLE-JP-AI-0002.1).

In the example 45, the relative clause is '*which is very interesting in a day'* and *the* relative pronoun '*which'* points back to the antecedent '*several different places'*. However, this antecedent '*several different places'* is plural one and the relative clause should grammatically follow it with a relative clause like '*which are very interesting in a day'*. Similar to this example, it is possible to notice the same error types in the examples 46, 47, 48, 49 and 50.

46. Now at university, I also read **books** *which* **is not** major (ICLE-JP-AI-0003.1).

47. And **government** *which* **pursue** only their own profit (ICLE-JP-KO-0010.1).

48. I read an article that says **the pets** *which* **is** thrown away is increasing (ICLE-JP-KO-0013.2).

49. They had to put on **clothes** *which* **looks like** the Yamato (ICLE-JP-SWU-0018.2).

50. You have chance to speak to English speaking people and reading **article** *which* **are** written by English (ICLE-JP-SWU-0027.3).

4.4.2. Errors of Relative Pronoun *Which* in Non-restrictive Relative Clauses

As both Turkish and Japanese have both inflectional features in constructing grammar structures, there is considerable similarity between two languages. Similar to the Turkish, Japanese have not got relative pronouns like English. Since there are not any relative pronoun or any other indication of the case relation between the modifying section and the N it modifies, it is sometimes only from the context or knowledge of the word that it can be determined. There is also no formal distinction between restrictive and non-restrictive relative clauses. However, in English, non-restrictive relative clauses are set off by comma intonation, and written between commas (Kaiser et al. 2001, p. 447). Accordingly, Japanese university students also tended to make errors in non-restrictive relative clauses by adding punctuation without determining whether the relative clauses with pronoun *which* has the function of adding extra or optional information about the head noun.

4.4.2.1. Punctuation Errors in the Relative Pronoun which in Non-restrictive Relative Clauses

51. For example **textbooks**, *which* pupils read, involve a lot of messages about gender (ICLE-JP-KY-0002.1).

As can be seen in the example 51, the non-restrictive relative clause is created by the learner such as '*which pupils read'* However, the relative clause in this example should not add extra or optional information about the antecedent '*textbooks*'. This antecedent needs to be defined and the relative clause in this example does not need to be written by punctuation. There are similar error types in the examples 52, 53 and 54.

52. Juice of cosmetics is dropped and **the process of changing eyes**, *which* become injured, is watched (ICLE-JP-KO-0024.2).

53. They play music with **traditional instrument**, *which* is made of bamboo (ICLE-JP-KO-0025.2).

54. Perhaps human being can obtain **some valuable things**, *which* can not be gained from man (ICLE-JP-KO-0027.2).

The antecedents in the following examples 55, 56, and 57 are all proper nouns. Thus, the learners should have regarded this relative clause as a non-restrictive relative clause.

55. This case's factor is **ANA** *which* is a sponsor of Flugels (ICLE-JP-SWU-0019.2).
56. And I also entered **STEPS Musical Company** *which* is Keio University's club (ICLE-JP-KO-0019.1).
57. The other day I went to watch a movie **The Mexican** *which* is performed by Brad Pitt and Julia Roberts (ICLE-JP-KO-0009.2).

4.4.2.2. Errors of Sentential Relative Clauses

It is a special type of non-restrictive relative clause. It does not point back to a noun phrase but to a whole clause or sentence. Among relative pronouns, only the relative pronoun *which* is used in sentential relative clauses (Leech & Svartvik, 2002, p. 387). Similar to Turkish, there is also no formal distinction between restrictive and non-restrictive relative clauses in Japanese. Thus, they also tend to show punctuation confusion in sentential relative clauses. In the examples 58, 59 and 60, the learners neglect adding comma in the sentential relative clause.

58. However, there are also many temptations in this campus *which* makes me to speak in Japanese (ICLE-JP-SWU-0002.4).
59. They would hurt the neighbors as they did before *which* might result in the ex-criminal's isolation from the society (ICLE-JP-TF-0033.1).
60. For instance, his neighbor may have prejudice against him *which* makes uneasy or annoying (ICLE-JP-TF-0037.1).

4.4.3. Choice of Relativizer pronoun

Unlike Turkish students' errors in the choice of relativizer pronoun, Japanese learners tend to combine the relative pronoun *which* with abstract nouns such as society, period, point. For instance, in the examples 61, 62 and 63, although the antecedent are abstract nouns, the learners tend to combine it with the relative pronoun *which* rather than the relative pronouns *in which* or *where*.

61. Now we live in the society *which* we can select our will freely (ICLE-JP-SWU-0005.2).

62. Because child have a period *which* could acquire first language (ICLE-JP-SWU-0012.2).

63. A learner remembers the rules and s/he will be able to find a point *which* needs to be careful by themselves (ICLE-JP-SWU-0002.3).

CHAPTER 5

CONCLUSION

5.1. Introduction

This study attempts to examine the use of relative clauses with the relative pronoun *which* in Turkish university students' written English as to underuse, overuse and error types and is to compare it with the use of native English writing, native Turkish writing and another learner data Japanese university students' written English. Furthermore, the study aims at reaching the possibility of whether Turkish university students' written language is affected by their mother tongue Turkish or only the feature of interlanguage.

5.2. Evaluation of Research Questions

1. Do Turkish university students use the relative pronoun *which* both in restrictive and non-restrictive relative clauses to the same extent as native speaker American university students?

Turkish and English have differences with respect to construction in the relative clauses. That is to say, Turkish does not involve the operation of specifying head noun with a relative pronoun. Considering this, Turkish university students were expected to have some difficulty in constructing the relative clauses with the relative pronoun *which* both in restrictive and non-restrictive relative clauses because Turkish does not have the relative pronouns and it does not have non-restrictive relative clause. The overall results of the study have indicated that there is a significant difference of the use of relative clauses with the relative pronoun *which*. In fact, Turkish learners use more the relative pronoun *which* in their essays than the native speaker American students (362 vs. 200 usages). This overall result indicates that Turkish university students tend to overuse the relative pronoun *which* in their written English.

On the other hand, there does not seem to be a significant difference between Turkish and American university students as regard to the use of non-restrictive relative clauses. While the number of occurrence might reveal similarity (72 vs. 74), the error

analysis of non-restrictive relative clauses with the relative pronoun *which* presents non-native features of Turkish university students. As a consequence, Turkish university students do not seem to use the relative pronoun *which* to the same extent native speaker American university students.

2. Does Turkish university students' use of the relative pronoun *which* both in restrictive and non-restrictive relative clauses give any evidence of transfer from their native Turkish?

Turkish learners overused the relative pronoun *which* in their essays than the native speaker American students (362 vs. 200 usages). To be more specifically, Turkish learners have tended to overuse the relative pronoun *which* in restrictive relative clauses (290 vs. 126 usages). Similarly, in TUC Turkish university students tended to use the structure more frequent than in the TICLE texts by Turkish learners of English (1597 vs. 362). It can be claimed that the cases of overuse seem to suggest transfer from native Turkish. Particularly, the use of non-restrictive clauses by Turkish learners seems to give an evidence of transfer from native Turkish since there is no formal distinction between restrictive and non-restrictive relative clauses (see section 2.7.2.1)

3. Do relative clauses with the relative pronoun *which* occur most frequently as restrictive or non-restrictive clauses in the texts?

Biber et al. (1999) examined that there are more occurrences of restrictive relative clauses than non-restrictive relative clauses in native English texts of academic prose. Similar to the result of the study by Biber et al. (1999), the results of the present study show that the relative pronoun *which* occur most frequently as restrictive clauses in the whole texts both learner and native corpora.

4. If there is a significant difference among underuse and overuse of the relative clauses with the relative pronoun '*which'* both in restrictive and non-restrictive relative clauses, does this stem from the mother tongue or is it interlanguage universal?

As a result of the analysis of the use of relative pronoun *which* in relative clauses, it can be stated that some noteworthy differences and similarities have been revealed. First, in comparison to the overall use of the relative pronoun '*which'* by Japanese learners, there seems a significant difference. Turkish learners have tended to overuse the structure than Japanese learners (362 vs. 311). However, both learner corpora analysis show that these two learner corpora present more occurrences of relative pronoun *which* than native speaker American corpus. For instance, Turkish university students vs. native speaker American students (362 vs. 200). Moreover, native speaker of Turkish university students overuse it than Turkish learners of English (1597 vs. 362). As Turkish has participle suffixes, the noun of the noun phrase that a relative clause modifies namely the head noun can have several different relationships with the relative clause. Thus, Turkish learners might tend to confuse the relationship between the head noun and the relative clauses. This might be the explanation of mother tongue influence for Turkish learners.

The errors produced in both learner corpora show similarity in terms of Chi-square statistics and error description in the data analysis. Namely, the results of analysis in both restrictive and non-restrictive relative clause errors do not reveal a significant difference. We have examined the quantitative analysis of the relative clauses with the relative pronoun *which* in Turkish university students' argumentative essay writing and compared it with the use in native speaker American English writing. We have also compared it with the use in learner corpus called JPICLE and native speaker Turkish corpus called TUC.

5.3. Implications for English Language Teaching

Data Driven Learning activities can be regarded as learner autonomy, ranging from teacher-led and relatively closed concordance-based exercises to entirely learner-centered corpus projects (Mukherjee, 2004, p. 10). In the process of both teaching and learning the relative clauses, teachers and students should be aware of real language use. More importantly, the ELT teachers should provide a reliable and extensive source of target structure. In addition such an extensive source can be analyzed both by teachers and by students. For example, individual learners can use a concordance display of their own mistakes as a starting point for data-driven learning activities. The examination of learner data by the learners themselves will motivate many more learners to reflect on

language use and thus raise foreign language awareness. With the help of reaching to KWIC (key word in context) concordance packages, online both native and learner corpora, the students can have chance of examining the relative clauses with the relative pronoun *which* about the usage, meaning, and the linguistic complexity they challenge such as punctuation in non-restrictive relative clauses, sentential relative clauses, and specific difficulties that Turkish students faced in restrictive relative clauses. For this very reason, considering the relative clauses with the relative pronoun *which,* a data-driven learning approach would be effective in improving the deficiencies occurred in the Turkish university students' written English.

5.4. Further Research

This study is only limited to the relative clauses with the relative pronoun *which* in written texts. For further research, the use of relative clauses with the relative pronoun *which* can be investigated in spoken corpora. Furthermore, studies which describe the use of all relative pronouns in relative clauses might provide a broader insight into what happens in the learner language. Specifically, as Turkish learners might also tend to prefer relative pronouns *that* instead of *which*, the future researches on the relative pronouns such as *that, in which* could give us a broader picture of mother tongue influence. Hence, we believe that corpus studies are important in terms of comparison and contrasting what non-native and native speakers of a language can do in comparable texts.

REFERENCES

Altenberg, B. & Tapper, M. (1998). The use of adverbial connectors in advanced Swedish learners' written English. In S. Granger (Ed.), *Learner English on computer* (pp. 80-93). New York: Longman.

Baker, P. Hardie, A. & McEnery, T. (2006). *A Glossary of corpus linguistics* Edinburgh: Edinburgh University Press Ltd.

Beaugrande, R. (2001). ‘‘If I were you...’’ Language standards and corpus data in EFL. *Revista Brasileira de Linguistica Aplicada*, 1, 116-117.

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

Biber, D. (1998). *Variation across speech and writing.* Cambridge: Cambridge University Press.

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

Biber, D. & Reppen R. (1998). Comparing native and learner perspectives on English grammar: a study of complement clauses. In S. Granger (Ed.), *Learner English on computer* (pp. 145-158). New York: Longman.

Can, C. (2010). İkinci dil edinimi çalışmalarında bilgisayar destekli bir Türk öğrenici İngilizcesi derlemi: ICLE’nin bir alt derlemi olarak TICLE. *Language Journal,* 144, 1-16.

Chafe, W. (1992). The importance of corpus linguistics to understanding the nature of language. In J. Svartvik (Ed.), *Directions in corpus linguistics. Proceedings of Nobel Symposium 82, Stockholm 4-8 August 1991* (pp. 79-92). Berlin: Mouton.

Charteris, B. J. (2004). *Corpus approaches to critical metaphor analysis.* Hants: Palgrave- MacMillan.

Chomsky, N. (1984). *Modular approaches to the study of the mind*. San Diego: San Diego University Press.

Cornilescu, A. (1981). Non-restrictive relative clauses: an essay in semantic description *Revue Roumaine de Linguistique*, 26(1), 41-67.

De Cock, S. (1998). A recurrent word combination approach to the study of formulae in the speech of native and non-native speakers of English. *International Journal of Corpus Linguistics*, 3 (1), 59-80.

Depraetere, I. (1995). Factors requiring, promoting and excluding the use of a (non-) restrictive relative clause. *Leuvense Bijdragenü*, 83(4), 433-458.

Fillmore, C. (1992). Corpus linguistics or computer-aided armchair linguistics. In J. Svartvik (Ed.), *Directions in corpus linguistics. Proceedings of Nobel Symposium 82, Stockholm 4-8 August 1991* (pp. 79-92). Berlin: Mouton.

Gabrielatos, C. (2005). Corpora and language teaching: Just a fling or wedding bells? *Teaching English as a Second Language* (pp. 1-35). Retrieved January 5, 2010 from, http://tesl-ej.org/ej32/a1.html

Gavioli, L. (2005). *Exploring corpora for ESP learning.* Amsterdam: John Benjamins.

Gavioli, L. (2001). *The learner as researcher: introducing corpus concordancing in the classroom*. Houston, TX: Athelstan.

Göksel, A. & Kerslake, C. (2005). *Turkish: A comprehensive grammar*. New York: Routledge.

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

Granger, S., Dagneaux, E., Meunier, F. & Paquot, M. (2009). *The international corpus of learner English. version 2 handbook and CD-ROM* (pp. 40-43). Louvain-la-Neuve: Presses Universitaires de Louvain. Retrieved June 1, 2011 from http://www.i6doc.com

Granger S., Dagneaux, E. & Meunier F. (2002). *The international corpus of learner English handbook and CD-ROM*. Louvain-la-Neuve: Presses Universitaires de Louvain

Greenbaum, S. & Quirk, R. (2003). *A student's grammar of the English language.* London: Longman.

Halliday, M. (1991). *Corpus studies and probabilistic grammar.* New York: Longman.

James, C. (1992). Awareness, consciousness and language contrast: New departures in contrastive linguistics. *Innsbrücker Beitrage zur Kulturwissenschaft*, 2, 79-92.

Johns, T. & King, P. (1991). *Classroom Concordancing*. Birmingham: University of Birmingham.

Kennedy, G. (1998). *An Introduction to corpus linguistics*. London: Longman.

Kaiser, S. Ichikawa, Y. Kobayashi N. & Hilofumi Y. (2001). *Japanese: A Comprehensive Grammar.* New York: Routledge.

Lado, R. (1957). *Linguistics across cultures*. Ann Arbor: University of Michigan.

Lam, P. & Hung, J. (1998). The use of multi-word verbs in advanced Chinese ESL learners. In S. Granger (Ed.), *Learner English on computer* (pp. 80-82). New York: Longman.

Leech, G. & Svartvik J. (2002). *A Communicative grammar of English.* London: Longman.

Lightbown, P. M. & Spada, N. (1999). *How languages are learned*. Oxford: Oxford University Press.

Lorenz, G. (1999). Overstatement in advanced learners' writing: stylistic aspects of adjectives intensification In Granger (Ed.), *Learner English on computer* (pp. 53-66). New York: Longman.

McCarthy, M. & O'Keeffe, A. (2010). Historical Perspective: What are Corpora and How Have They Evolved? In O'Keeffe, A. and McCarthy, M. (Eds) *The Routledge handbook of corpus linguistics*. London: Routledge

McDavid, V. (1977). Which in relative clauses. *American Speech*, 52 (2), 76-83.

McEnery, T. & Wilson, A. (2001). *Corpus linguistics.* Edinburgh: Edinburgh University Press.

Meyer, C. F., (2002) *English corpus linguistics*. Cambridge: Cambridge University Press.

Meyer, C. F., (2004). *English corpus linguistics: An Introduction.* Cambridge: Cambridge University Press.

Mukherjee, J. (2004). The state of the art in corpus linguistics: three book length perspectives. *English Language and Linguistics*, 8 (1), 103-119.

Mukherjee, J. & J. Rohrbach (2006). Rethinking applied corpus linguistics from a language-pedagogical perspective: New departures in learner corpus research. In Ketteman, B. and Marko, G. (Eds), *Planning and Gluing Corpora: Inside the Applied Corpus Linguist's Workshop*. (pp. 205-232). Frankfurt am Main: Peter Lang.

Nesselhauf, N. (2004). Learner corpora and their potential for language teaching. In Sinclair (Ed.), *How to Use Corpora in Language Teaching* (pp. 125-152). Amsterdam: John Benjamins,

Quirk, R., Greenbaum, S. Leech, G. & Svartvik, J. (1985). *A comprehensive grammar of the English language.* London: Longman.

Saussure, F. de (1966). *Course in general linguistics*. New York: McGraw- Hill.

Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics*. 10(3), 209-231.

Sinclair, J. (Ed.). (1990). *Collins COBUILD English grammar.* London: Harper Collins.

Stubbs, M. (2001). *Words and phrases: Corpus studies of lexical*. London: Blackwell Publishers.

Tao, H. & McCarthy, M. J. (2001). Understanding non-restrictive which-clauses in spoken English, which is not an easy thing. *Language Sciences*, 23(6), 651-677.

Tono, Y. (2003). Learner Corpora: design, development and applications. *Proceedings of the Corpus Linguistics Journal,* 5(3), 800-809.

Virtanen, T. (1998). Direct questions in argumentative student writing. In Granger (Ed.), *Learner English on computer* (pp. 94-118). New York: Longman.

Virtanen, T. (1998). Argumentative uses of the progressive in NS and NNS student compositions: notes on clause status and grounding. In Granger (Ed.), *Learner English on computer* (pp. 119-120). New York: Longman.

Wikipedia. (2011). Text Corpus. Retrieved May, 12 from http://en.wikipedia.org/wiki/Text_corpus

Yamashita, J. (1994). An analysis of relative clauses in the Lancaster/IBM spoken English corpus. *English Studies,* (75) 1, 73-84.

CURRICULUM VITAE

PERSONAL DETAILS

Name: Mustafa ÇOBAN
Date of Birth: 16.01.1987
Place of Birth: Nazilli
Telephone: 05348442763
E-mail: muscoban09@gmail.com

EDUCATION

2009 – 2011 : MA at Çukurova University Institute of Social Sciences
English Language Teaching Department
Adana, Turkey

2005 - 2009 : BA at Çukurova University Faculty of Education
English Language Teaching Department
Adana, Turkey

2008 – 2009 : Nazareth College University
Leadership and American Culture Education
Rochester, New York, the USA
(Fulbright / State Department Scholar)

2007 – 2008 : Antwerp University – Artesis Hogeschool
Faculty of Education Antwerp, Belgium
(Erasmus Scholar –Fall Semester)

2001 - 2005 : Nazilli Super High School
Foreign Language Department
Aydin, Turkey

JOB EXPERIENCES

2011 - present : EFL Instructor at Çağ University Department of Preparation School Mersin, Turkey

2010 - 2011 **:** EFL Instructor at Uzman School of English Adana, Turkey

2009 - 2010 **:** EFL Instructor at American Cultural Association Language School Adana, Turkey

CONFERENCES ATTENDED

Date	**Institution**	**Title**
5-7 Mayıs	Çukurova Üniversitesi	25. Ulusal Dilbilim Kurultayı (Poster presented)
26-28 May	University of Silesia, Poland	The Affective Dimension in Language Teaching and Learning (Paper presented)