

GAZİ UNIVERSITY
INSTITUTE OF EDUCATIONAL SCIENCES
DEPARTMENT OF FOREIGN LANGUAGES TEACHING
ENGLISH LANGUAGE TEACHING PROGRAM

DISCOURSE MARKERS AND SPOKEN ENGLISH:
NONNATIVE USE IN THE TURKISH EFL SETTING

PHD DISSERTATION

BY
Asuman AŞIK

Ankara
June, 2012

GAZİ UNIVERSITY
INSTITUTE OF EDUCATIONAL SCIENCES
DEPARTMENT OF FOREIGN LANGUAGES TEACHING
ENGLISH LANGUAGE TEACHING PROGRAM

DISCOURSE MARKERS AND SPOKEN ENGLISH:
NONNATIVE USE IN THE TURKISH EFL SETTING

PHD DISSERTATION

Asuman AŞIK

Supervisor: Assoc. Prof. Dr. Paşa Tefik CEPHE

**Ankara
June, 2012**

JÜRİ ONAYI

Asuman AŞIK'ın “**Discourse Markers and Spoken English: Nonnative Use in the Turkish EFL Setting**” başlıklı tezi 05.07.2012 tarihinde, jürimiz tarafından İngiliz Dili Eğitimi Ana Bilim Dalında Doktora Tezi olarak kabul edilmiştir.

Adı Soyadıİmza

Başkan:

Üye :

Üye (Tez Danışmanı):

Üye :

Üye :

ACKNOWLEDGEMENTS

Within this long journey in writing a PhD thesis, there have been many people to whom I send my heartfelt thanks. First and foremost, I am mostly indebted to my supervisor, Assoc. Prof. Paşa Tervik CEPHE for his guidance and valuable comments all the way through. When I feel lost in so many details of the thesis, he always helped me clear my mind on the issues so I was able to move on.

I would like to also express my deepest gratitude to the thesis committee members, Prof. Dr. Abd÷lvahit ÇAKIR and Prof. Dr. Gülsev PAKKAN who supported me with their careful insights and believed in my thinking.

I would like to thank also to my colleagues in the department; Asst. Prof. Dr. H. Hande UYSAL and Dr. Kemal Sinan ÖZMEN for their careful reviews on my thesis and Asst. Prof. Dr. Korkut Uluç İŞİSAĞ for helping me in recording sessions. My special thanks also go to my precious friends who believed in me, motivated me with their support and cheered me up when I needed.

I am also thankful to my parents and my sisters who helped me when I always needed them. Finally, I am also indebted to my beloved husband, Erhan AŞIK, who encouraged me all the time and supported me both professionally and emotionally; and my daughter, Yağmur AŞIK, who always livened me up with her love and jolliness.

ABSTRACT**DISCOURSE MARKERS AND SPOKEN ENGLISH:
NONNATIVE USE IN THE TURKISH EFL SETTING**

AŞIK, Asuman

PhD Dissertation, English Language Teaching Program

Supervisor: Assoc. Prof. Dr. Paşa Tevfik CEPHE

June-2011, 200 pages

This dissertation aims at identifying the discourse markers used by Turkish nonnative speakers of English, their occurrences in their spoken English discourse by comparing them with the ones used in native speakers' spoken discourse and their use of functions. For these purposes, the study used both quantitative and qualitative research methods. Within the quantitative side of the study, a research corpus was composed with the course presentations of the Turkish undergraduate students studying at Gazi University throughout transcription process. To compare the data, transcripts of student presentations from University of Michigan with the help of MICASE Corpus were also attained. The occurrences of the discourse markers in both corpus were determined. The results show that Turkish nonnative speakers of English have a lack of variety in using discourse markers in their spoken English and use discourse markers in a limited number. As for the qualitative side of the study, the functions of the discourse markers were given examples from two corpus, which revealed that nonnative speakers do not benefit from the variety of functions of the discourse markers in spoken discourse. Along with these findings, the study highlights the importance of the need for awareness-raising of Turkish nonnative speakers in using discourse markers in their spoken English discourse and recommends considerable implications for English language teaching.

Key Words: Discourse Analysis, Spoken Discourse, Nonnative Spoken Discourse, Discourse Markers, Corpus Linguistics

ÖZET

SÖYLEM BELİRLEYİCİLERİ VE KONUŞMA İNGİLİZCESİ: İNGİLİZCE’NİN YABANCI DİL OLDUĞU TÜRK ORTAMINDA KULLANIMI

AŞIK, Asuman

Doktora, İngiliz Dili Eğitimi Ana Bilim Dalı
Tez Danışmanı: Doç. Dr. Paşa Tefik CEPHE
Haziran-2012, 200 sayfa

Bu doktora çalışması, anadili İngilizce’den farklı olan Türk konuşucuları tarafından kullanılan söylem belirleyicilerini, anadili İngilizce olan konuşucuların sözlü söylemlerinde kullanılanlarla karşılaştırarak, konuşma İngilizce söylemindeki tekrar sıklıklarını ve işlevlerinin kullanımını araştırmayı amaçlamaktadır. Bu amaçlar için, araştırma, hem nitel hem de nicel araştırma yöntemlerini kullanmıştır. Araştırmanın nitel yanı içerisinde, çevri yazı yöntemiyle, Gazi Üniversitesi’nde okuyan Türk lisans öğrencilerinin ders sunumlarını içeren bir araştırma bütüncesi oluşturulmuştur. Verileri, karşılaştırmak için, MICASE bütüncesi yardımıyla ile Michigan Üniversitesi’ndeki öğrenci sunumlarının yazılı metinlerine erişilmiştir. Her iki bütüncedeki söylem belirleyicilerinin tekrar sıklığı belirlenmiştir. Sonuçlar, Türk İngilizce konuşucularının konuşma İngilizcesi’nde söylem belirleyici çeşitliği eksikliğine sahip olduklarını ve söylem belirleyicilerini sınırlı sayıda kullandıklarını göstermektedir. Çalışmanın nicel yönü içinse, söylem belirleyicilerinin işlevleri, her iki bütünceden örneklendirilmiştir; bu da, anadili İngilizce olmayan konuşucuların, sözlü söylemde söylem belirleyicilerinin işlev çeşitliğinden yararlanmadıklarını ortaya çıkarmıştır. Bu bulgularla birlikte, araştırma, konuşma İngilizcesi söyleminde söylem belirleyicilerinin kullanımı konusunda Türk konuşucularda farkındalığı artırma ihtiyacının önemini vurgulamaktadır ve İngilizce dil eğitimi için dikkate değer çıkarımlar önermektedir.

Anahtar Kelimeler: Söylem Analizi, Sözlü Söylem, Ana Dili Olmayanların Sözlü Söylemi, Söylem Belirleyicileri, Bütünce Dilbilimi

TABLE OF CONTENTS

JÜRİ ONAYI.....	i
ACKNOWLEDGEMENTS	ii
ABSTRACT.....	iii
ÖZET	iv
TABLE OF CONTENTS	v
LIST OF TABLES	viii
LIST OF FIGURES	ix
CHAPTER 1:INTRODUCTION.....	1
1.1. Introduction.....	1
1.2. Statement of the Problem.....	3
1.3. Purpose of the Study	4
1.4. Scope of the Study	5
1.5. Methodology	5
1.6. Significance of the Study	7
1.7. Limitations of the Study	8
1.8. Definition of Key Concepts	8
CHAPTER 2: REVIEW OF LITERATURE.....	10
2.1. Introduction.....	10
2.2. Discourse Analysis	10
2.2.1. Historical Overview.....	10
2.2.2. Definition.....	11
2.2.3. The Scope of Discourse Analysis.....	12
2.2.4. Corpus Linguistics.....	14
2.2.5. Discourse Analysis and Language Teaching.....	15
2.3. Discourse Markers	16
2.3.1 Terminology of Discourse Markers.....	17
2.3.2. Classification of Discourse Markers	21

2.3.3. Characteristics of Discourse Markers.....	25
2.3.3.1 Connectivity.....	26
2.3.3.2 Multi-functionality/Polyfunctionality	27
2.3.3.3 Optionality	28
2.3.3.4. Non-truth conditionality	28
2.3.3.5 Weak clause association	29
2.3.3.6 Initiality.....	30
2.3.3.7 Orality	30
2.3.3.8 Multi-categoriality	31
2.4. Theories Underlying Discourse Markers	32
2.4.1. Coherence-based Account of Discourse Markers	33
2.4.2. Relevance-based Account of Discourse Markers.....	42
2.5. Discourse Markers and Language Teaching.....	44
CHAPTER 3: METHODOLOGY	47
3.1. Introduction.....	47
3.2. Corpus Selection	47
3.3. Data Collection	48
3.3.1. Participants	49
3.3.2. Setting.....	49
3.3.3. Instruments and Procedures for Data Collection.....	49
3.3.3.1. Recording.....	50
3.3.3.2. Transcription.....	51
3.4. Corpus Description	52
3.5. Data Analysis	54
CHAPTER 4: ANALYSIS OF RESULTS.....	57
4.1. Introduction.....	57
4.2. Identifying Discourse Markers and Non-discourse Markers in Research Corpus and MICASE Corpus	57
4.3. Findings of Analysis in the Research Corpus	61
4.4. Results of Analysis in MICASE Corpus.....	69
4.5. Comparative Analysis.....	76

4.6. Analysis of Functions of Discourse Markers in the Research Corpus and MICASE Corpus.....	83
4.6.1 Interpersonal Category	84
4.6.2. Referential Category.....	88
4.6.3. Structural Category.....	90
4.6.4. Cognitive Category.....	92
4.7. Discussion	95
CHAPTER 5: CONCLUSION	100
5.1. Introduction.....	100
5.2. Summary	100
5.3. Pedagogical Implications	103
5.4. Further Research	105
REFERENCES.....	107
APPENDICES.....	118
APPENDIX A: CLASSROOM OBSERVATION SHEET.....	118
APPENDIX B: MICASE TRANSCRIPTION AND MARK-UP CONVENTIONS	119
APPENDIX C: SAMPLE TRANSCRIPTS FROM RESEARCH CORPUS.....	123
APPENDIX D: SAMPLE TRANSCRIPTS FROM MICASE CORPUS	140
APPENDIX E: SAMPLE ACTIVITIES	185

LIST OF TABLES

Table 1: Pragmatic functions of discourse markers	23
Table 2: Planes of talk on which markers function	35
Table 3: Description of MICASE transcripts	53
Table 4: Description of transcripts of the research corpus	53
Table 5: Occurences of Discourse Markers in the first ten transcripts in the research corpus.....	62
Table 6: Occurences of discourse markers in the second ten transcripts in the research corpus.....	63
Table 7: Occurences and frequencies of the discourse markers in the research corpus .	65
Table 8: Occurences of discourse markers in MICASE transcripts	70
Table 9: Occurences and frequencies of the discourse markers in MICASE corpus	72
Table 10: Frequency level of the discourse markers according to the median of the list within the research corpus	78
Table 11: Frequency level of the discourse markers according to the median of the list within MICASE corpus	79
Table 12: Comparative results of the frequencies in the research corpus and MICASE corpus.....	81
Table 13: A core functional paradigm of discourse markers in pedagogic discourse	83

LIST OF FIGURES

Figure 1: A discourse model.....	34
Figure 2: First 20 most frequent discourse markers in the research corpus.....	69
Figure 3 : First 20 most frequent discourse markers in MICASE corpus.....	76

CHAPTER 1

INTRODUCTION

1.1. Introduction

In the current era, learning a second or a foreign language has a paramount role in composing a good communication in cross-cultural environments. Within this atmosphere, English, as being *lingua franca*, has become the language for the majority of people all around the world in terms of communication for commerce, trade, education and research, so there has been a need for learning and teaching English to fulfill several purposes. This necessity leads authors to a myriad of methods, approaches or theories about how English should be taught and learned. Starting from the 1950s, there have been several methods, approaches and designs suggested and applied in language classrooms (Gattegno, 1972; Curran, 1976; Johnson and Paulston, 1976; Wilkins, 1976; Terrell, 1977; Krashen, 1981; Finocchiaro and Brumfit, 1983).

During the last decades, the importance of communicative competence in foreign language learning and teaching has gained utmost significance as the ability to use language to communicate effectively is a fundamental issue. In parallel with communicative competence, pragmatics and particularly discourse analysis have become an inevitable part of foreign language learning and teaching as these are the fields interested in the relationship between language and the contexts in which it is used, how people really use the language, which is a very striking opposition to studying artificially created domains. In particular, discourse analysis tries to find out the relationship between form and function by analyzing any language in use through written texts or spoken data so discourse analysis has a pivotal role for language learners' pragmatic and communicative competence.

Moreover, Trillo (2002) states that “native and non-native speakers of language have different linguistic development which consists of two tracks: formal track and pragmatic track” (p.770). Native speakers of a language develop both tracks simultaneously by means of natural language contact although non-native learners of a

language develop formal and pragmatic tracks through formal instruction. However, it is not easy to implement pragmatic track through educational syllabuses. Thus, foreign language learners use certain forms inappropriate to the context and the setting. By time, these certain forms can be fossilized. Trillo (2002) calls this process “pragmatic fossilization” which is defined as “the phenomenon by which a non-native speaker systematically uses certain forms inappropriately at the pragmatic level of communication” (p.770).

Learners of foreign language need authentic tasks or materials to be competent communicatively. Discourse markers, which are very significant within the field of discourse analysis, are essential in teaching English communicatively. Discourse markers are expressions such as those in bold in the following sentences:

- a. A: I like him. B: **So**, you think you’ll ask him out then.
- b. John can’t go. **And** Mary can’t go either.
- c. Will you go? **Furthermore**, will you represent the class there?
- d. Sue left very late. **But** she arrived on time.
- e. I think it will fly. **After all**, we built it right. (Fraser, 1999:931).

Although throughout literature, some other terms such as “discourse particles”, “connectives”, “pragmatic expressions” or “pragmatic markers” are used by some researchers to define the expressions illustrated above, the term “discourse markers” (DMs) is more commonly preferred and employed by researchers analysing English discourse.

Fraser (1990) has likened the effect of discourse markers to that of “discourse glue” (p. 385) as they unite the utterances within discourse. Discourse markers are defined by Schiffrin (1987) as “sequentially dependent elements which bracket units of talk” (p. 31). Moreover, discourse markers are units of talk that can be used for several purposes, as in the following:

- to initiate discourse,
- to mark a boundary in discourse (shift/partial shift in topic),

- to preface a response or a reaction,
- to serve as a filler or delaying tactic,
- to aid the speaker in holding the floor,
- to effect an interaction or sharing between speaker and hearer,
- to bracket the discourse either cataphorically or anaphorically,
- to mark either foregrounded or backgrounded information. (Müller, 2005:9)

As it can be concluded, discourse markers have considerable importance in teaching English since they contribute a lot to the pragmatic and communicative competence of speakers. Svartvik (1980) illustrates this essentiality as in the following:

If a foreign language learner says *five sheeps* or *he goed*, he can be corrected by practically every native speaker. If, on the other hand, he omits a *well*, the likely reaction will be that he is dogmatic, impolite, boring, awkward to talk to etc, but a native speaker cannot pinpoint an ‘error’. (p.171)

Regarding Svartvik’s example, an utterance or a sentence that lacks discourse markers cannot be labelled as ungrammatical. However, the hearer or the reader may find the speaker or the author boring, routinised or impolite. Müller (2005) also points out that “if we take it for granted that discourse markers have such a decisive role to play in native speaker communication as the authors of discourse marker analyses claim, then we must assume that they are important elements to be learned by non-native speakers as well” (p. 14).

1.2. Statement of the Problem

As pragmatic competence which is defined by the ability to communicate effectively and involves the knowledge beyond the level of grammar by Thomas (1983), it should definitely be integrated in learning and teaching a foreign language. Teaching pragmatic competence has been searched by many authors (Bardovi-Harlig, 1996; Rose and Kasper, 2001; Bardovi-Harlig and Griffin, 2005). Furthermore, Crozet (2003) states that some of the rules that govern interactions but that are not immediately obvious have been referred to as invisible rules. Within these invisible rules, discourse analysis has gained fundamental importance to help learners attain pragmatic competence. Native speakers of English apply these invisible rules without noticing what kind of elements

they should include to their discourse. Particularly, in spoken discourse, native speakers use naturally certain units of talk. Hence, discourse markers are among these units of talk uttered by the speakers to make their speech more understandable and rich; as Crystal comments (1988), they serve as the “oil which helps us perform the complex task of spontaneous speech production and interaction smoothly and efficiently” (p. 48), so they are highly important in teaching English to foreign language learners.

In addition, Fung and Carter (2007) propose that language learners should learn discourse markers “in order to facilitate more successful overall language use and at the very least for reception purposes” (p. 434). Like many other non-native speakers of English, Turkish non-native speakers also have difficulties in enriching their conversations in English although they may be much more competent in using textual or structural coordinates in written discourse in English. Thus, it is necessary to find out whether Turkish non-native speakers of English use these particular discourse items, namely discourse markers, adequately and effectively in their spoken discourse and what the level of frequencies of discourse markers is in their speech or whether they benefit from these markers to make their speech more understandable, rich, polite or colorful. Therefore, the study will allow for a better understanding of language use and acquisition that occurs as part of these students’ spoken English discourse.

1.3. Purpose of the Study

Due to the significance of discourse markers in spoken discourse of native speakers of English, there is a necessity to investigate these specific discourse items in spoken discourse of nonnative speakers of English. Thus, the study, as an overall purpose, aims at identifying the discourse markers used by Turkish nonnative speakers of English. For this purpose, the study tries to find the answers to the following research questions which are the driving force of the current study:

- Which discourse markers are used by Turkish non-native speakers of English in spoken discourse?
- What is the frequency level of the discourse markers used by Turkish non-native speakers of English in spoken discourse?

- Are there any differences between the discourse markers used by Turkish non-native speakers and native speakers of English in spoken discourse?
- What are the prevailing functions of the discourse markers employed by Turkish non-native speakers when compared with the ones of native speakers according to four categories (interpersonal, referential, structural and cognitive)?

1.4. Scope of the Study

The study focuses mainly on the discourse markers used by Turkish non-native speakers of English. To reach the objectives of the study, undergraduate students of Department of English Language Teaching (ELT) at Gazi University, Turkey were taken as the representative sample group for Turkish non-native speakers of English. The Department of English Language Teaching of Gazi University is one of the leading and populous departments within the field in Turkey with more than 1250 bachelors and nearly 40 graduate students.

The study intended to reach its objectives by investigating twenty student presentations done by twenty senior-undergraduate students of this ELT Department with upper level proficiency in English who have completed the courses related to English skills and the main theoretical basis of language teaching.

Moreover, for the comparative dimension of the research, the study also focuses on the discourse markers used by the senior-undergraduate students of University of Michigan, USA. Similarly, the student presentations of University of Michigan are taken as the sample group.

1.5. Methodology

As the research aims at identifying the discourse markers that Turkish non-native speakers of English use in their spoken discourse and compare the results with the ones uttered by native speakers of English, the study is based on a corpus-driven approach. Thus, the methodology and the analysis of the study includes the essential characteristics that a corpus-based study should have, particularly highlighted by Biber, Conrad and Reppen (1998) as in the following:

- It is empirical, analyzing the actual patterns of use in natural text;
- It utilizes a large and principled collection of natural texts, known as a “corpus”, as the basis for the analysis;
- It makes extensive use of computer for analysis, using both automatic and interactive techniques;
- It depends on both quantitative and qualitative analytical techniques. (p. 4)

This study focuses on two types of corpus taken as the basis of the analysis which are the corpus of Turkish non-native speakers and the corpus of native speakers of English. The latter one, called MICASE, is taken from Michigan Corpus of Academic Spoken English which is a spoken language corpus available on-line. This corpus was taken as the sample to constitute the second corpus, which is called as the research corpus. The research corpus was constituted for the study by the researcher.

In order to realize the stated objectives of the study, the data were collected through audio recordings of student presentations of the students studying at Gazi University English Language Teaching Department. Transcription method was chosen as transcripts of classroom recordings provide an excellent record of “naturally occurring interaction” (Silverman, 1993). Therefore, audio recordings of utterances made by students were transcribed in standard orthography and all transcript lines containing the linguistic item in question, whether discourse marker or not, was extracted and sorted.

The study is based on mixed method data analysis; quantitative and qualitative analysis. The quantitative method emphasizing the descriptive value for spoken discourse of recurrent patterns and of frequency distribution was used to reach the frequency counts of discourse markers used by Turkish non-native speakers. Moreover, a qualitative analysis including the study of core functional paradigm of discourse markers in academic discourse based on the multi-categorial model proposed by Fung and Carter (2007) was also conducted. The four categories which are interpersonal, referential, structural and cognitive were taken as the functional paradigm so that the illustrations from the corpora were presented to show how discourse markers were used.

1.6. Significance of the Study

According to the recent analyses of corpora of spoken interaction, discourse markers are among the top ten word forms (Allwood, 1996 cited in Carter and Fung, 2007); so throughout history, there have been numerous studies about discourse markers in English (Svartvik, 1980; Östman, 1981; Schiffrin, 1986; Aijmer, 1987; Schourup, 1985; Erman, 1987). Moreover, the studies of discourse markers in other languages have also been conducted by many authors (Bazzanella, 1990; Gupta, 1995; Chen and He, 2001). However, the studies about the use of discourse markers in English by second or foreign language speakers are limited. Hays (1992), Trillo (1997), Müller (2004) and Fung and Carter (2007) are notable authors within this field of investigation who searched the use of discourse markers by different groups of speakers that are originally using another language.

As English takes place among the education system with a pivotal role in Turkey, like many other countries, it is necessary to search for the characteristics of spoken discourse of Turkish nonnative speakers of English. Moreover, there has not been any research done about this specific subject, particularly, discourse markers used by Turkish-nonative speakers of English according to the review conducted by the researcher. Thus, this study is significant in order to identify the discourse markers of Turkish non-native speakers so as to provide essential implications for teaching these units of talk to language learners and to make them gain discourse-pragmatic competence in English.

Moreover, this study is also significant in order to provide comparative analysis between native speakers and Turkish non-native speakers of English in using these particular discourse elements so as to enrich the field of discourse analysis like the similar studies stated above.

1.7. Limitations of the Study

The study has certain limitations which are put forward by the nature of discourse analysis. As the study focuses on spoken discourse, it may have some drawbacks when compared to the analyses done in written discourse. In particular, within this study, data collection procedures and transcription process of student presentations to compose the research corpus are highly arduous and time-consuming. Thus, the number of the presentations for the research corpus is limited to twenty. Moreover, the research corpus is limited to only the presentations of students studying in ELT Department of Gazi University while the corpus of native speakers is limited to only four transcripts of nearly 18 presentations of the students studying in University of Michigan.

Another limitation is about the analysis of the functions of the discourse markers. The study focuses on dealing with the occurrences of discourse markers while it is limited to only giving illustrations for the functions of several instances of discourse markers. Due to the multifunctional characteristic of discourse markers, it is not possible to analyse each function in each instance of discourse markers.

1.8. Definition of Key Concepts

Discourse Analysis:

The analysis of discourse is, necessarily, the analysis of language in use. As such, it cannot be restricted to the description of linguistic forms independent of the purposes or functions which these forms are designed to serve in human affairs (Brown & Yule, 1983: 1).

Corpus:

Corpus, plural corpora, is a collection of linguistic data, either compiled as written texts or as a transcription of recorded speech. The main purpose of a corpus is to verify a hypothesis about language - for example, to determine how the usage of a particular sound, word, or syntactic construction varies (Crystal, 1992).

Corpus Linguistics:

Corpus linguistics studies the principles and practice of using corpora in language study. The main focus of corpus linguistics is to discover patterns of authentic language use through analysis of actual usage (Krieger, 2003).

CHAPTER 2

REVIEW OF LITERATURE

2.1. Introduction

This chapter presents the detailed background to the study by pointing out the general framework of the study; particularly discourse analysis, description and characteristics of discourse markers and previous researches about discourse markers.

2.2. Discourse Analysis

2.2.1. Historical Overview

In 1960s and early 1970s, the importance of communicative competence gained fundamental importance in second/foreign language teaching. The emergence of communicative competence did not happen at once and only. There have been many simultaneous and interactive theories, studies, researches done, all of which argue the relationship between language and context. All of these studies are related to each other. Thus, it is not possible to claim that only one of them is the most important one and has influenced the others. The studies done about *pragmatics*, *discourse analysis*, *text linguistics*, *conversational implicatures* seem to be discussing similar arguments and each of them mainly focuses on a particular aspect of the relationship between language and context.

Discourse analysis is generally described as the study of language in use, both in the form of written texts or spoken data. It is fundamentally related to several disciplines such as linguistics, semiotics, psychology, anthropology and sociology in the 1960s and early 1970s. Although it is claimed that the emergence of discourse analysis was in 1960s, its history dates back to 1952 when Zellig Haris published a paper titled as “Discourse Analysis” which included the studies about the links between the text and its social situation. Then, Dell Hymes was the one who interested in the study of speech

in its social setting from a sociological perspective in the 1960s. Austin (1962), Searle (1969) and Grice (1975) followed Dell Hymes with their studies upon the study of language as social action, speech act theory, conversational maxims and pragmatics. Moreover, M.A.K. Halliday's functional approach to language and focus on text linguistics also stressed the importance of social functions of language and contributed greatly to the future studies upon discourse analysis.

2.2.2. Definition

Discourse analysis is such a vast and ambiguous field that it is strenuous to define and delimit its borders as the analysis of the study. Brown and Yule (1983) defines discourse analysis as in the following:

The analysis of discourse, is necessarily, the analysis of language in use. As such, it cannot be restricted to the description of the linguistic forms independent of the purposes or functions which these forms are designed to serve in human affairs. (p. 1)

Brown and Yule (1983) claim that linguistic forms should not be analysed independently from their purposes or functions in real interactions. McCarthy (1996) also develops this argument by stating that "discourse analysis is not entirely separate from the study of grammar and phonology but discourse analysts are interested in a lot more than linguistic forms" (p. 8). He also emphasized the preoccupations of discourse analysts overlap in an important sense with the approach to communicative language teaching which focuses on the functions or speech acts that pieces of language perform. Furthermore, Stubb (1983) presents what kinds of forms that discourse analysis aims to study as in the following:

Discourse analysis attempts to study the organization of language above the sentence or above the clause, and therefore to study larger linguistic units, such as conversational exchanges or written texts. It follows that discourse analysis is also concerned with language in use in social contexts, and in particular with interaction or dialogue between speakers. (p. 1)

As Stubb points out above, the scope of the discourse analysis includes linguistic units in their widest sense and they are not limited to a sentence or a clause. The aim is

to analyse beyond the sentence/sentences or utterances and how they are related in interactions.

Schiffrin (1987) also lists the assumptions of discourse analysis as in the following:

1. *Language always occurs in a context:* As for the “context” in the assumption, Schiffrin explains that it includes not only cognitive contexts in which past experience and knowledge is stored and drawn upon but also social contexts through people can draw upon institutional and interactional orders to construct definitions of situation and action.

2. *Language is context sensitive:* Language is potentially sensitive to all of the contexts in which it occurs, and, even more strongly, that language reflects those contexts because it helps to constitute them.

3. *Language is always communicative:* Language is always communicative either because it is directed toward a recipient (immediate or eventual), because it is intended to be so directed, and/or because it is attended by a recipient.

4. *Language is designed for communication.* The primary purpose of language is to communicate. Thus, language is in a change constantly in order to fulfill the needs of communication. (p. 4-6)

2.2.3. The Scope of Discourse Analysis

The scope of discourse analysis is so vast that it includes several examples of utterances. Discourse analysis not only deals with the description and analysis of spoken interaction but also written and printed words like “newspaper articles, letters, stories, recipes, instructions, notices, comics, billboards, leaflets” (McCarthy, 1996:12). Since the written texts are easy to be specified, spoken texts may need more attention in terms of description, scope and focus of analysis. McCarthy (1996) lists some different types of speech as in the following:

Telephone calls (business and private)
 Service encounters (shops, ticket offices, etc.)
 Interviews (jobs, journalistic, in official settings)
 Classroom (classes, seminars, lectures, tutorials)
 Rituals (sermons, weddings)
 Monologues (speeches, stories, jokes)
 Language-in-action (talk accompanying doing: fixing, cooking, assembling, demonstrating, etc.)
 Casual conversation (strangers, friends, intimates)
 Organising and directing people (work, home, in the street) (p. 119)

Discourse analysts are interested in both written forms and spoken forms of language. However, the analysis of each form differs from another. Moreover, to gather data in spoken interaction is a difficult task to do in terms of statistics of the distribution of different types of speech in people's daily lives. Siniajeva (2005) discusses what kind of features of written discourse differ from spoken language by stating the following:

Linguistically, written text tends to be more complex, with longer sentences, more complex clauses, greater information load, with the higher number of lexical or content words per clause. Unlike spoken interaction, in written discourse there is no common situation: the situation has to be inferred from the text. The words themselves must carry all the shades of meaning which, in spoken discourse, could be conveyed by non-verbal behavior. (p. 11)

However, there are some advantageous issues of written texts over spoken data. Written texts allow writer the possibility of editing the text or the reader to read back and forth and the statements in written texts are generally grammatically correct whereas spontaneous conversations may not always be well formed.

Moreover, McCarthy (1996) points out that "both types of discourse are dependent on their immediate contexts to a greater or lesser degree" (p.149). Written texts may require particular shared knowledge while spoken discourse is supported with intonation and actions.

2.2.4. Corpus Linguistics

In parallel with the developments in discourse analysis, there has emerged a new impetus to discourse analysis, which is corpus linguistics. As a brief definition, corpus linguistics is the study of collecting, structuring and analysing a large amount of discourse (corpus), with the help of computers. It allows through the use of computerized technology certain operations such as “quantifying (counting the number of given words or sentences), concordancing (producing lists of linguistic items and their immediate linguistic context in order to determine syntactic, semantic and pragmatic properties) and parsing (seperating sentences into grammatical parts)” (Fortuno, 2006:58).

Although there is not a consensus about what the corpus linguistics is, as Taylor (2008) highlights, “whether it is a tool, a method, a methodology, a methodological approach, a discipline, a theory, a theoretical approach, a paradigm (theoretical or methodological), or a combination of these” (p. 180), corpus linguistics has gained its place within the literature.

Historical background of corpus linguistics dates back to Randolp Quirk (1959) who organized a corpus of both spoken and written British English, named as the Survey of English Usage (SEU) Corpus and then by Nelson Francis and Henry Kučera, the Brown Corpus was composed as a sample of printed American English in 1961. With the experience gained after SEU and Brown Corpus, Jan Svartvik, who is another pioneer of the field, made a survey of Spoken English in 1975 and the result was London-Lund Corpus (LLC).

As stated by Leech (1991), “within the thirty years since 1961, corpus linguistics has gradually extended its scope and influence and become a mainstream in itself” (p. 9). With the advancements in computer technology, the number of different types of corpora and publications about them have increased. This prolific progress was also stated as a ‘second generation’ (Leech, 1991) which includes John Sinclair’s Birmingham Collection of English Text and the Longman/Lancaster English Language Corpus in 1980s. Other comprehensive projects in 1990s is called COBUILD that covers 450 million words of spoken and written British English and British National

Corpus (BEC) with 100 million word collection of samples, which are followed by Cambridge International Corpus (CIC) which hold 600 million words in both British and American English discourse. Moreover, MICASE (Michigan Corpus of Spoken Academic English) which is used for the current research and stated in detail in *Chapter 3: Methodology*, is another latest corpus that serves grave functions for the insights in spoken English discourse.

2.2.5. Discourse Analysis and Language Teaching

It may not be inarguably accepted that the insights of discourse analysis are applicable in several and definable ways to language teaching although the main focus of discourse analysis is not to present analysis of texts to be used in language teaching. The detailed studies of discourse are not only used for the sake of linguistics but also for the sake of language teachers as the analysis of written or spoken texts both contribute a lot to make many systematic ways of language teaching clearer (McCarthy and Carter, 1995; McCarthy, 1996; Celce-Murcia and Olshtain, 2000; Hunston, 2002).

The analysis of spoken interaction is highly fundamental to assist second/foreign language learners gain native-like fluency. Since the spoken data are collected from real-life conversations and interactions, they will suggest authentic examples for the non-native learners. McCarthy (1996) also supports the contribution of discourse analysis to language teaching as “discourse analysis can supply data where intuition cannot be expected to encompass the rich detail and patterning of natural talk” (p. 145). McCarthy (1996) exemplifies this claim as in the following:

Teachers will make up their own minds as to whether their methods and techniques need rethinking in the light of what discourse analysts say, but, as with all new trends in linguistic theory and description, it is important that discourse analysis be subjected not only to the scrutiny of applied linguists but also to the testing grounds of practical materials and classroom activities. (p. 171)

McCarthy (1996) also claims that discourse analysis is not a method for teaching languages, and moreover it does not claim to be a methodology. However, it provides analysis of the linguistic forms in their widest sense and context-sensitive, which is

much more different from traditional analysis of the linguistic forms such as lexis, grammar and phonology. It is definitely impossible to say that lexis, grammar and phonology are not useful in linguistics and language teaching. However, they are not adequate to provide learners competence in a second/foreign language. Thus the findings of discourse analysis might be applied as complementary to the ones of grammar, lexis or phonology.

Moreover, with the help of corpora and its tools, corpus linguistics also provide a more objective analysis of language patterns which can be applied to discourse analysis and language teaching. According to Krieger (2003), “a corpus-based analysis can investigate almost any language patterns--lexical, structural, lexico-grammatical, discourse, phonological, morphological--often with very specific agendas such as discovering male versus female usage of tag questions or children's acquisition of irregular past participles”. Thus corpus linguistics offer several ways of implementation to language teaching ranging from syllabus design to materials development.

2.3. Discourse Markers

The study of discourse markers is rooted in discourse analysis. In the last twenty years, the interest towards discourse markers increased a lot. Many authors analysed discourse markers from different points of view. Numerous researches done about how the use of discourse markers contribute to pragmatic and communicative competence of speakers. The diversity of researches upon discourse markers made difficult to delimit the terminology, the characteristics or the classification of the discourse markers, which is also pointed out by Schourup (1999) below:

While it is widely agreed that such expressions play a variety of important roles in utterance interpretation, there is disagreement in regard to such fundamental issues as how the discourse marker class should be delimited, whether the items in question comprise a unified grammatical category, what type of meaning they express, and the sense in which such expressions may be said to relate elements of discourse. (p. 227)

The study of discourse markers began in 1980s with many scholars studying these items simultaneously as they were mostly found necessary components to be included

in written or spoken discourse. Discourse markers as a subject of study were first mentioned by Levinson (1983), but only briefly. The first comprehensive analysis of these linguistic units was conducted by Schiffrin (1987). She suggested the general framework of discourse markers and analyzed specifically the linguistic items such as *and, because, but, I mean, now, oh, or, so, then, well* and *y'know* through a sociolinguistic research which was carried out in unstructured conversations and spontaneous speech. At the same time Blakemore was also studying upon discourse markers from a relevance theoretic approach. Then Fraser, Schourup, Aijmer, Saez and others followed them; which is presented in detail in the following.

2.3.1 Terminology of Discourse Markers

Throughout history, several researchers such as Schiffrin (1987), Blakemore (1987), Halliday and Hasan (1992), Fraser (1993), Andersen (2001), Aijmer (2002), Trujillo Saez (2003) and so forth have differently label the phenomenon “discourse marker”. The terms given by many researches are *pragmatic markers* (Fraser, 1999), *discourse markers* (Schiffrin, 1987), *discourse particles* (Schourup, 1985), *discourse connectives* (Blakemore, 1987), *cue phrases* (Knott and Dale, 1994), *interactional signals*, *pragmatic expressions* and so on. Thus, there has been a terminological problem within the field. However, each term has a particular characteristic that makes it different from others and every linguist who labeled the specific term has his/her own justification. To be specific, throughout the study, the term “discourse marker” has been used.

As is stated above, the first brief introduction about today’s discourse markers was given by Levinson (1983) as in the following:

... there are many words and phrases in English, and no doubt most languages, that indicate the relationship between an utterance and the prior discourse... It is generally conceded that such words have at least a component of meaning that resists truth-conditional treatment... What they seem to do is indicate, often in very complex ways, just how the utterance that contains them is a response to, a continuation of, some portion of the prior discourse. (p. 87-88)

As Levinson states above, in languages there are mostly words or phrases which provide a kind of contextual bridge between the prior discourse and the following discourse. Meanwhile, Quirk, Greenbaum, Leech and Svartvik (1985) emphasise the interactional effect of these words and their importance to develop an ongoing and intimate relationship with people by explaining that the phrases such as *well*, *y'know*, *really* are 'sharing devices' and 'intimacy signals' in everyday conversation:

It is easily demonstrable that these play, from the point view of grammatical structure, no part in transmission of information, yet only is our present-day colloquy constantly embellished with them, but popular talk stretching back to Shakespeare and beyond has been similarly peppered with these apparently useless and meaningless items...since the desire to feel that the hearer is sharing something with one seems to be fundamental in the urge to speak, these sharing devices, these intimacy signals in our everyday talk, are of considerable importance. (p. 178-79)

Schiffrin (1987) proposed that discourse markers could be looked from a "more theoretical level as members of a functional class of verbal (and non-verbal) devices which provide contextual coordinates for ongoing talk" (p.41) and pointed out an operational definition by describing "*discourse markers as sequentially dependent elements which bracket units of talk*" (p. 31). She describes in detail why she prefers to use these terms to define discourse markers. First, she uses *units of talk* as a more general term rather than sentence, proposition, speech act as the term *units of talk* transcend the sentence due to the optionality of discourse marker within a sentence. The discourse markers can be just one phrase, one word or multi-word and they are not restricted to any particular position in the sentence structure. Especially, in conversations, discourse markers are used in several positions. Secondly, she uses the term *brackets* as generally discourse markers are either anaphoric or cataphoric devices, such as the following examples respectively:

- a. He came home late *y'know*.
- b. *Y'know* he came home late. (Coll, 2009:50)

The significance of *sequentially dependent* within the definition refers that discourse markers do not depend on the units of talk of which the discourse is composed, but the discourse as a whole. Thus, discourse markers are defined as they

integrate forms, meanings and actions to make overall sense out of what is said, which finally contributes to discourse coherence. She exemplifies that *oh, well, and, but, or, so, because, now, then, I mean, y'know* are discourse markers and proposes that these markers mainly serve three functions:

1. they act as contextual coordinates for utterances by locating them on one or more planes of discourse;
2. they index adjacent utterances to the speaker, the hearer, or both;
3. they indicate the utterance to prior and/or subsequent discourse. (p. 19)

Moreover, Fraser (1999) focused on specifically “*What are DMs? What are not DMs? What is the grammatical status of DMs? And what do DMs link?*” and provided a comprehensive definition of DMs stated below by referring to the relationship between the utterances that follow each other as *S1* for the prior utterance and *S2* for the following one:

A class of lexical expressions drawn primarily from the syntactic classes of conjunctions, adverbs and prepositional phrases. With certain exceptions, they signal a relationship between the interpretation of the segment they introduce *S2*, and the prior segment, *S1*. They have core meaning which is procedural, not conceptual, and their more specific interpretation is ‘negotiated’ by the context, both linguistic and conceptual. (p. 931)

However, Andersen (2001) prefers to use the term “pragmatic marker” instead of discourse marker as “the label ‘pragmatic’ means to suggest a relatively low degree of lexical specificity and a high degree of context-sensitivity” (p. 40) and pragmatic markers have textual function which contributes a relationship between the current message and the previous message. Moreover, Lenk (1997) tries to differentiate both terms as in the following:

Studies that investigate pragmatic markers often focus more on the interactional aspects between the participants that are expressed through the use of particles. One of the most prominent functions of discourse markers, however, is to signal the kinds of relations a speaker perceives between different part of the discourse. (p. 2)

Another suggestion about the term was made by Blakemore (1987) who discussed some discourse markers like *and, after all, you see, but, moreover, furthermore and so* and called them “discourse connectives” by stating that these expressions “constrain the interpretation of the utterances that contain them by virtue of the inferential connections they express” (p. 105).

Moreover, Schourup (1999:229) was among the ones who preferred to use the term discourse marker instead of discourse particle as ‘particle’ is related to syntax while discourse markers generally signal a functional class consisting of items that belong to several syntactic classes. Furthermore, he also states that discourse particle tends to be used much more inclusively than discourse marker. Thus, the term ‘discourse marker’ has a narrower range and has been subject to more precise attempts at definition.

Hansen (1998) also defines discourse markers as “linguistic items which fulfill a non-propositional, metadiscursive (primarily connective) function, and whose scope is inherently variable, such that they may comprise both sub-sentential and supra-sentential units” (p. 236) and discusses that “semantically, markers are best seen as processing instructions intended to aid the hearer in integrating the unit hosting the marker into a coherent mental representation of the unfolding discourse” (p. 236).

The study of discourse markers are seen as an inevitable part of pragmatics so they should not only be studied grammatically. Aijmer (2002) supports this claim by defining discourse markers as a “class of words with unique formal, functional and pragmatic properties” (p.2). He also states that discourse markers are difficult to analyse grammatically and their literal meanings are ‘overridden’ by pragmatic functions involving the speaker’s relationship to the hearer, to the utterance or to the whole text. Furthermore, Aijmer (2002) states the necessity of discourse markers in utterance interpretation as in the following:

Discourse particles seem to be dispensable elements functioning as signposts in the communication facilitating the hearer’s interpretation of the utterance on the basis of various contextual cues. This does not mean that discourse particles are meaningless decorations or a verbal ‘crutch’ in discourse indicating a lack of speaker proficiency,

but they are better dealt with in pragmatics or in discourse analysis than in semantics.
(p. 2)

As is stated above, discourse markers should be studied within pragmatics and discourse analysis instead of semantics as they provide listeners have better understanding and interpretation of the speaker's utterances and serve contextual cues for the interpretation. Although discourse markers are stated as dispensable units in communication, there is not a consensus upon its terminology and description.

2.3.2. Classification of Discourse Markers

Apart from the terminological problem of discourse markers, there is also the problem of classification and what kind of words are accepted as discourse markers. This is one of the main controversial issue which Fraser (1999) also accepts by stating that "researchers have agreed that DMs are lexical expressions that relate discourse segments, but they have disagreed on how they are defined and what functions they carry" (p.931). There has been dispute over whether discourse markers should be classified according to their syntactic groups or their functions. Moreover, Schourup (1999) claims that "even when an item is widely accepted as a DM, there can be disagreement about which instances of the item qualify" (p. 241) and he gives examples from Schiffrin and Redeker who admit *I mean* and *y'know* as DMs but Redeker excludes literal uses of these expressions (e.g. *you know what Hasidic is?*) as DMs while Schiffrin accepts.

To support this issue, Jucker (1993) claims that "there is no generally accepted list of discourse markers in English" (p. 436). Thus, while there seems to be general agreement for some elements, other elements are of more doubtful status (*because, and, then* are included by Schiffrin (1987) but not Schourup (1982), while *hey* and *aha* are included by Schourup but not by Schiffrin).

Moreover, Fraser (1999) defines DMs as a pragmatic class, lexical expressions drawn from the syntactic classes of conjunctions, adverbials, and prepositional phrases. He treats DMs as a subclass of pragmatic markers. Schourup (1999) summarizes

Fraser's four types of pragmatic markers corresponding to four distinct message types as in the following:

1. Basic markers: These markers specify the force of the basic message, which is the message which refers to sentence proposition.

For example: *Admittedly*, I was taken in.

Admittedly in the above sentence is a basic marker as it signals that the proposition uttered by *I was taken*.

2. Commentary markers: These markers express a comment on the basic message.

For example: *Stupidly*, Sara didn't fax the correct form in on time.

Stupidly signals the speaker's comment on Sara's failure to send the fax. These kind of markers include markers of assessment (*sadly*), manner-speaking (*frankly*), emphasis (*mark my words*), mitigation (*if you don't mind*) and consequent-effect (*to sum up*), evidential markers (*certainly*), and hearsay markers (*allegedly*).

3. Parallel markers: these kind of markers signal a message additional to the basic message.

For example: *Waiter*, please bring me another fork.

Get your *damned* shoes off the table.

In sum, Fraser defines discourse markers as a separate fourth type of pragmatic marker which contributes nothing to truth-conditionality and serves connectivity.

As the categorization has diversity, there has been dispute over what kind of linguistic units can be accepted as discourse markers. To give an example, Lee-Goldman's (2010) research about "*No* presents that although *yeah* is one of the most frequently used discourse marker which has several functions such as agreement and acknowledgement, topic management and speaker shift, little attention has been paid to *no*" (p.1), thus proposes *No* as a discourse marker that has particular functions such as topic shift, misunderstanding management and turn-taking conflict resolution. Lee-

Goldman (2010) points out that *no* can function as marker on the basis findings of Schegloff (2001) and Schegloff (1992), provided as in the following:

- 1 Roger To tell you the truth, I'd rath- I'd, I'd – would like
 2 to avoid more than one I_C_S_I meeting per day, if possible.
 3 [((laugh)) But – ((laugh))] =
 4 Brian [O_K.
 5 Roger = I mean. I don't know. Whatever.
 6 Brian -- > No, that's fine. (cited in Lee-Goldman, 2010:1-2)

Brinton (1996) claims that discourse markers serve a variety of pragmatic functions and classifies DMs according to their functions in two categories; *textual functions and interpersonal functions*. The category of textual functions include the meanings that the speaker structures as text by creating cohesive passages of discourse. In other words, those functions are related to the context. Second category, interpersonal functions, are related to social exchange, in a way, the role of the speaker and the role assigned to the hearer. Castro (2009) adopted the inventory functions devised by Brinton in the following table and the examples given in the table are the functions of DMs used by the participants in the specific class sample of Castro's study:

Table 1: Pragmatic functions of discourse markers

Textual Functions	To initiate discourse, including claiming the attention of the hearer	Opening frame marker	<i>so; ok; now</i>
	To close discourse	Closing frame marker	<i>ok; right; well</i>
	To aid the speaker in acquiring or relinquishing the floor.	Turn takers. (Turn givers)	<i>um; eh; and</i>
	To serve as a filler or delaying tactic used to sustain discourse or hold the floor.	Fillers	<i>ok; well; now</i>
	To indicate a new topic or a partial shift in topic.	Turn keepers Topic switchers	<i>and; because; so</i>
	To denote either new or old information	Information indicators.	<i>so; and; and then; because</i>
	To mark sequential dependence.	Sequence/relevance markers	<i>well; I mean, you know; like</i>
	To repair one's own or	Repair markers.	<i>well; I mean, you</i>

	others' discourse.		<i>know; like</i>
Interpersonal functions	Subjectively, to express a response or a reaction to the preceding discourse including also back-channel signals of understanding and continued attention while another speaker is having his/her turn.	Response/reaction markers	<i>yeah; oh; ah; but; oh yeah; well; eh; oh really?</i>
		Back-channel signals	<i>mhm; uh huh; yeah</i>
	Interpersonally, to effect cooperation or sharing, including confirming shared assumptions, checking or expressing understanding, requesting confirmation, expressing difference or saving face (politeness).	Cooperation, agreement marker	<i>ok; yes; yeah; mhm</i>
		Disagreement marker	<i>but; no</i>
		Checking understanding markers	<i>ah; I know; yeah; mhm; yes</i>
		Confirmation-seekers	
Face-savers			

(Castro, 2009)

Another classification is done by Redeker (1990) who divides discourse markers into two categories: those that mark *ideational structure*, such as connectives and temporal adverbials (e.g. *and*, *meanwhile*, or *now*) and those which mark *pragmatic structure* (e.g., *oh*, *alright*, or *well*). She finds that, to some extent, the use of DMs of these two types are complementary. According to her research, ideational markers are used more than pragmatic markers in all cases, but the number of ideational markers used goes down when speakers use a high number of pragmatic markers.

Jucker and Smith (1998) examined differential use of DMs based on the relationship between interlocutors and divided DMs into reception markers (e.g. *oh*, *yeah*, and *okay*) and presentation markers (e.g. *like*, *you know* and *well*). They found that the presentation markers like *well* and *you know* were used more in interactions between friends, and the reception markers *oh* and *yeah* were used more between strangers.

Fung and Carter (2007) have categorized DMs within four categories, which is taken as the basis for qualitative analysis of the research. They arrange their categories by following Maschler (1994, 1998) and point out that one discourse marker can function more than only one specific one. Their multifunctional category includes four categories which are *interpersonal*, *referential*, *structural* and *cognitive category*. Within interpersonal category, DMs are used to signal shared knowledge (such as *you know*, *you see*, *see*, *listen*) and to indicate responses like agreement, confirmation and acknowledgement (such as *Okay*, *oh*, *right/alright*, *yes*, *I see*, *great*, *oh great*, *sure*). These kind of markers are used to indicate the attitudes of the speaker (such as *well*, *I think*, *you know*, *sort/kind of*, *like*, *just*, *to be frank*, *etc.*). Another category is referential category which includes DMs that are on a textual level and used to mark relationships between verbal activities preceding by conjunctions: cause (*because/cos*), consequence (*so*), contrast (*but*, *and*, *yet*, *however*, *nevertheless*), coordination (*and*), disjunction (*or*), digression (*anyway*) and comparison (*likewise*, *similarly*). The third category is structural category in which DMs are used to indicate the discourse in progress. Signposting opening and closing of topics (*now*, *OK*, *right*, *well*, *by the way*, *let's start*, *let me conclude the discussion*), indicating sequential relationships (*first*, *firstly*, *second*, *next*, *then*, *finally*) and marking topic shifts (*so*, *now and what about*, *how about*) are the DMs within this category. Another category is cognitive category including DMs that provide information about the cognitive state of speakers. These kind of DMs are used to instruct a mental representation of the discourse. The speaker uses these kind of DMs to denote the thinking process (*well*, *I think*, *I see*, *and*), reformulate (*I mean*, *that is*, *in other words*), elaborate (*like*, *I mean*), mark hesitation (*well*, *sort of*) and assess the listener's knowledge about the utterances (*you know*).

Briefly, it can be stated that the complex form of discourse markers in terminology also influence their classification. That's why several types of classifications are valid within the field.

2.3.3. Characteristics of Discourse Markers

Discourse markers have specific features that make them different from other phrases or clauses. Different theoretic backgrounds approach discourse markers in different points of view and thus this makes the characteristics of discourse markers

versatile. Jucker (1993) stated Hölker's (1991) list of four basic features of discourse markers as in the following:

1. they do not affect the truth conditions of an utterance.
2. they do not add anything to the propositional content of an utterance.
3. they are related to the speech situation and not to the situation talked about.
4. they have an emotive, expressive function rather than a referential, denotative, or cognitive function. (p.436)

Moreover, Schourup (1999) is the one who gathers the general characteristics of discourse markers according to connectivity, multifunctionality, optionality, non-truth conditionality, weak clause association, initiality, orality and multicategoriality.

2.3.3.1 Connectivity

Connectivity is one of the basic characteristic of DMs as discourse markers are used to establish a relationship between the current utterance and the previous utterance. However, different authors approach this characteristic of connectivity in different ways. As Schiffrin (1987) and Fraser (1988) are on the side of coherence-based approach while analysing discourse markers, they point out that DMs relate two textual units by contributing to inter-utterance coherence. On the other hand, Blakemore, who is on the side of relevance-based approach, states that it is better to view certain DMs not as necessarily relating two segments of text, but as relating the propositional content expressed by the current utterance to assumption that may or may not have been communicated by a prior utterance, as in the following example:

[Seeing someone return home with parcels]

So you've spent all your money. (Blakemore, 1987:86)

In the example above, there is not a sentence uttered before. Here *so* refers to a proposition derived from "observation of a state of affairs", that is, to a context in a wider sense. Thus, Blakemore differs from Schiffrin and Fraser's account regarding that it is not necessary to have two relevant textual units to use a discourse marker, there may have some other non-verbal signs to use the discourse marker. This kind of dispute over connectivity is highly relevant to two approaches about DMs; coherence-based

models of discourse and relevance approach. Moreover, Schourup (1999) discusses this characteristic as in the following:

If connectivity, however formulated, is considered criterial for DM status, it can be used to distinguish DMs from various other initial elements, such as illocutionary adverbials (frankly, confidentially), attitudinal adverbials (fortunately, sadly), and from primary interjections (yipes, oops); however, connectivity alone is insufficient to distinguish DMs from coordinators joining intrasentential elements. (p.231)

As Schourup (1999) also points out above, the characteristic of connectivity of discourse markers is not an adequate feature to call a linguistic form as a discourse marker. Furthermore, this characteristic is distinctive in the sense that makes discourse markers distinguishable from other initial elements.

2.3.3.2 Multi-functionality/Polyfunctionality

It has been concluded from the relevant researches that discourse markers are used to fulfill several functions. To give an example, *but* functions differently in two sentences:

John likes football; *but* Mary likes basketball.

John is a lawyer; *but* he is honest. (Coll, 2009:48)

Within the first sentence, *but* is used to indicate contrast while in the other sentence it is used to show denial of expectations. Moreover, Jucker (1993) claims that this polyfunctionality of discourse markers make it difficult to compose a unified description of them and then proposes three solutions for this problem.

The first solution is to say that a particular discourse marker, for instance *well*, is ambiguous and requires several separate entries in a lexicon. The second solution is to say that all uses can be related to one core meaning. The third solution, finally, does not accept the polyfunctionality but claims that – properly understood – all uses can be summarised under one general description. (Hölker, 1991:86, cited in Jucker, 1993:437)

Although the solutions above do not suggest a plausible explanation, it is better

to approach this feature of multifunctionality of discourse marker in a different way. To accept that discourse markers function in different ways and to analyse the issue particularly within the immediate context can shed light.

2.3.3.3 Optionality

There is another characteristic of DMs which is their *optionality* as syntactically and semantically. Fraser (1998) states that removal of a DM does not alter the grammaticality of its host sentence (p.22). For instance:

The others are going to Stoke. *However*, I am going to Paris.

The others are going to Stoke. I am going to Paris.

Schourup (1999), with the example above, points out that “if a DM is omitted, the relationship it signals is still available to the hearer, though explicitly cued” (p. 231). However, this does not mean that this characteristic of DM makes DMs useless or redundant. Behind this optionality lies that DMs guide the hearer to a particular interpretation and at the same time ruling out unintended interpretations. For example, in the above sentences, although it is not necessary to use *however* to signal the meaning, adding *however* to the utterance will ‘reinforce’ or ‘clue’ the interpretation intended by the speaker.

Schiffrin (1987) also supports this claim by stating that discourse markers are independent of sentential structure although markers often precede sentences, i.e. syntactic configurations of an independent clause plus all clauses dependent on it. Thus removal of a marker from its sentence initial position leaves the sentence structure intact. Furthermore, several markers – *y’know*, *I mean*, *oh*, *like* – can occur quite freely within a sentence at locations which are very difficult to define syntactically. Especially, in spoken discourse, many discourse markers can be placed in different locations, and also can be omitted without any change in syntax.

2.3.3.4. Non-truth conditionality

When analysed semantically, it is generally thought that DMs do not contribute anything to truth-conditions of the proposition expressed by an utterance. This

characteristic of DMs makes DMs different from ‘content’ words such as adverbial uses of words like *sadly*. Jucker (1993) exemplifies this characteristic as in the following:

A: but who has to buy it

B: *well* the – the state has to buy it but ... (p.436)

He also states that the presence or absence of *well* in B’s utterance does not change the truth-conditions and it does not add anything to the propositional content of B’s utterance. The utterance would be a true (or false) representation of B’s opinion in exactly the same circumstances if B omitted *well*. It does not add any information to the proposition that ‘the state has to buy it’. However, there are cases where *well* does affect the truth conditions and adds to the propositional content, such as a manner adverb (*She draws well*); a degree word (*You know that perfectly well*), a noun (*Everyone digs their own well*). In these cases, *well* is not a discourse marker.

However, this may not always be the case with all markers and all their uses. Coll (2009) provides examples for these cases:

(1) A. He was really tired. *However*, the noise did not let him sleep.

B. He was really tired. The noise did not let him sleep.

(2) A. John went to Paris and *therefore*, Mary went to Rome.

B. John went to Paris and Mary went to Rome. (p.46)

In the first example, the discourse marker *however* does not affect the truth conditions of either the previous sentence or the sentence it appears in. However, in the second example, the removal of *therefore* in 2b may be also the removal of contribution of the discourse marker to the truth conditions of the utterance.

2.3.3.5 Weak clause association

As discourse markers can be treated out of the syntactic structure or not a strong component within sentential structure, these linguistic forms are considered as having weak clause association. Schourup (1999) states that “weak clause association is frequently correlated with phonological independence” (p. 233) and she explains that “

might be true for many DMs, such as conjuncts and disjuncts in general, regardless of whether they occur within the clause or at its extremes” (p.233). However, lack of intonational integration may not be a necessary characteristic of DMs.

2.3.3.6 Initiality

When analysed syntactically, many authors such as Hansen (1997) and Schifffrin (1987) point out that “DMs generally introduce the discourse segments they mark”. General tendency of DMs in terms of their position in the utterance is to take place in initial position, such as in the following:

(After all/Now/However), corgis are an intelligent breed.

Corgis, *(after all/now/however)*, are an intelligent breed.

Well, now, you know, but that’s not what it says in the instruction manual.

Through the examples above, Schourup (1999) states that “the tendency of DMs to appear initially is probably related to their ‘superordinate’ use to restrict the contextual interpretation of an utterance: in general it will make communicative sense to restrict contexts early before interpretation can run astray” (p.233). However, there are many cases where DMs take place in medial or final position of an utterance, such as in the following:

She likes all kinds of music classical er mainly classical **I think**.

But ah since it’s for children, this can’t be too high the price, **I mean**.

He send his regards **actually**. (Fung and Carter, 2007:413)

As a resut, it can be pointed out that the position of discourse markers in spoken discourse can vary. The examples above show that different types of discourse markers can be used in initial, medial or final position of utterances for specific purposes.

2.3.3.7 Orality

Schourup (1999) claims that “some of DMs occur often in spoken discourse such as *well, by the way* while some are mostly found in written discourse such as *consequently, moreover*” (p.234). However, this characteristic can be relevant to the formality/informality of the DM or the utterance. The use of specific DMs may change

according to the modes of texts or style (written/discourse and formal/informal) . Moreover, Schourup (1999) explains this characteristic as the meaning of a marker may also ally it to one channel or the other by giving an example as in the following:

Some putative DMs such as *conversely* and *in contrast* encode a high degree of utterance planning (compare such ‘impromptu’ speech-linked DMs as *before I forget* and *by the way*); other DMs may be associated with speech because their meaning presupposes a familiarity with the addressee not typical of impersonally addressed writing. (p. 234)

In other words, in spontaneous speech, it may not be possible to organize the speech with conjunctives as they are used in written texts. Due to the very essence of the differences between written texts and spoken texts, the use of particular discourse markers in particular organizations may be different from each other.

Taboada (2006) also presents that “although there are some exceptions, most studies are conducted about the relations signalled by explicit discourse markers, particularly in written texts and non-linguistic signals such as *mood, modality or intonation* have been ignored in these studies” (p.573). Moreover, to give an example, she also cited Louwerse and Mitchell’s (2003) research. This study concludes that there are 10 times as many discourse markers in spoken as in written discourse and also when compared to discourse type according to informality and formality, informal discourse includes twice as many discourse markers as in formal discourse.

2.3.3.8 Multi-categoriality

The categorization of DMs is a difficult task to do as authors dealing with DMs discuss whether they should be classified in a syntactic or functional category. Schourup (1999) states that “DM status is independent of syntactic categorization: an item retains its non-DM syntactic categorization but does ‘extra duty’ as a non-truth-conditional connective loosely associated with clause structure” (p.234) and he also states the general categories including extrinsic DM functions such as adverbs (*e.g. now, actually, anyway*), coordinating and subordinating conjunctions (*e.g. and, but, because*), interjections (*e.g. oh, gosh, boy*), verbs (*e.g. say, look, see*), and clauses (*e.g. you see, I mean, you know*). Therefore, discourse markers are items that cannot be easily

categorized into groups as their functions and their lexemes can be different from each other.

2.4. Theories Underlying Discourse Markers

As discourse markers have been studied by several researchers, different approaches have been proposed to present an account upon discourse markers. There are two prominent approaches developed throughout literature: coherence-based account and relevance-based account. The researchers who developed coherence-based account of discourse markers are Schiffrin (1987), Fraser (1988, 1990), Redeker (1990, 1991), Zwicky (1985) and Giora (1997, 1998) as cited in Hussein (2009). The ones who are on the side of relevance-based account based their approach upon Sperber and Wilson's (1995) Relevance Theory. This group includes Diane Blakemore (1987, 1992, 2002), Regina Blass (1990) and Corrine Iten (1998).

On the very nature of both approaches, they discuss how the use of DMs contributes to discourse interpretation. Schourup (1999) states briefly that "coherence group argue that DMs play a major role in the interpretation of the text by signalling 'coherence' relations between discourse units" (p.240). In other words, the interpretation of a text depends on the identification of coherence relations between the units of that text. On the other hand, the researchers in the relevance group consider DMs as indicators or procedures in two dimensions. According to Blakemore (2000), first dimension is that they limit the inferential phase of utterance by guiding the interpretation and the second dimension is that they offer clues to enable the hearer/reader to notice the intended meaning without many effort. Briefly, the first group deals with discourse markers within 'coherence relations' while the second group studies DMs within 'pragmatic relations'.

In fact, it seems that both of them are similar but 'relevance' considers discourse as a cognitive entity while 'coherence' refers to a linguistic concept. However, within each group, there is disagreement about the semantic, pragmatic and structural status of DMs.

2.4.1. Coherence-based Account of Discourse Markers

The first assumptions about the discourse coherence was made by Halliday and Hasan (1976). They analysed textuality in detail, and proposed particularly two characteristics of texts which are *coherence* and *cohesion*. Both features are highly related to each other in the sense that coherence is an umbrella term which includes cohesion.

Cohesion is the term used to describe the structural, grammatical and lexical means by which sentences and paragraphs in the texts are linked and relationships between them established. Halliday and Hasan (1976) define cohesion “as the set of possibilities that exist in the language for making text hang together” (p.18). It is a semantic relation between an element in the text and some other element that is crucial to the interpretation of it, which is expressed partly through the grammar and partly through the vocabulary.

Halliday and Hasan (1976) point out that particularly in English, “the basic means of establishing cohesion is through the use of pronouns, determiners, conjunctions, conjuncts and adverbials to substitute, repeat, refer or omit items across a text” (p.4). Cohesion occurs where the interpretation of some element in the discourse is dependent on that of another. The one presupposes the other, in the sense that it cannot be effectively decoded except by shift to it. When this happens, a relation of cohesion is set up, and the two elements, the presupposing and the presupposed, are thereby at least potentially integrated into a text. For example:

They think ***so***.

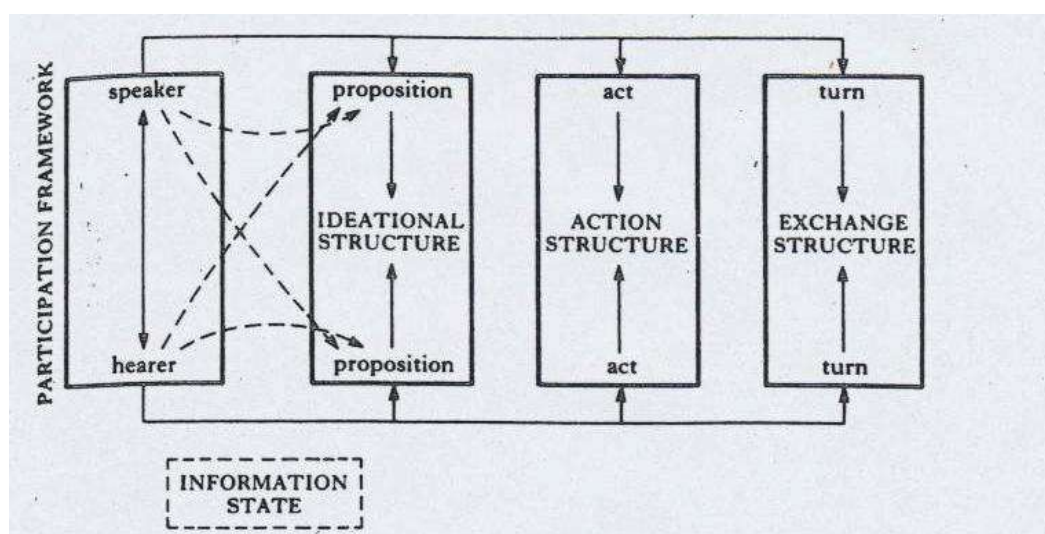
The sentence above could not be interpreted alone. The reader or the listener have to search for the referents to ***they*** and ***so***. Siniajeva (2005) states that “formal links between sentences and between clauses are known as cohesive devices, single instances of which are called ties” (p.18). Trujillo Saex (2001) separates different types of cohesive ties, such as: reference, substitution, ellipsis, lexical cohesion, conjunction and discourse markers. Therefore, Halliday and Hasan (1976:4) explains the concept of cohesion in detail such as in the following:

The concept of cohesion is a semantic one; it refers to relations of meaning that exist within the text and that define it as a text. Cohesion occurs where the interpretation of some element in the discourse is dependent on that of another. The one presupposes the other, in the sense that it cannot be effectively decoded except by resource to it. When this happens, a relation of cohesion is set up, and the two elements, the presupposing and the presupposed, are thereby at least potentially integrated into a text. (p.4)

Another important element of textuality is 'coherence' which is the logical relationships between the concepts. Cohesion contributes greatly to discourse coherence in terms of structural unity. Halliday and Hasan (1976) state briefly that the interpretation of a text should be analyzed within this feature of coherence. Similarly, Schiffrin (1987) proposes that DMs should be considered as linguistic devices that link adjacent units of talk to make the whole discourse coherent. Schiffrin analysed the DMs used by ordinary speakers by recording the interviews through tape-recorder. Thus, she composed a sociolinguistic corpus which includes the data of long transcribed speech units taken from the interviews.

Shiffirin (1987) proposes a model of coherence in talk, a model of discourse. The model focuses on local coherence, i.e. coherence that is constructed through relations between adjacent units in discourse. Schiffrin based her analysis of discourse markers on this model, as the chart displays in the following:

Figure 1: A discourse model



(Schiffirin, 1987:25)

Moreover, Schiffrin (1987) explains her discourse model as in the following:

My discourse model has both non-linguistic structures (*exchange and actions*) and linguistic structures (*ideational*). Speaker and hearer are related to each other, and to their utterances, in a participation framework. Their knowledge and meta-knowledge about ideas is organised and managed in an information state. Local coherence in discourse is thus defined as the outcome of joint efforts from interactants to integrate knowledge, meaning, saying and doing. (p.29)

This model consists of five planes of talk: *information state*, *participation framework*, *ideational structure*, *action structure* and *exchange structure*. Exchange structure is the result from the alternations between participants and their relation to each other, so the units of talk in an exchange structure are called turns. Moreover, the structures in action plane are defined by acts that occur in constrained linear sequences and are interpreted in the order of their occurrence. In addition, in the ideational structure, the units are semantic, that is, ideas. In addition to these three structures, there are two more structures which are participation framework, which reflects the way in which participants can be related to their utterances, briefly, includes speaker/hearer relations and speaker/utterance relations and information state which involves the organization and management of knowledge and meta-knowledge of the participants. The following table summarizes the possible effects that discourse markers have in the five planes of talk.

Table 2: Planes of talk on which markers function

Information state	Participation framework	Ideational structure	Action structure	Exchange structure
oh*	oh		oh	
well	well*	well	well	well
		and*	and	and
		but*	but	but
		or*		or
so	so	so*	so	so
because		because*	because	
	now	now*		

then		then*	then	
I mean	I mean*	I mean		
y'know*	y'know	y'know		y'know

(*) Primary function.

(Schiffrin, 1987:316)

She exemplifies her model by stating that DMs such as *and*, *but*, *or*, *so* and *because* are operative on the ideational structure. Such markers can indicate three types of relations that contribute to the configuration of idea structures: cohesive relations, topic relations and functional relations. On the other hand, the DMs such as *well*, *oh*, *now*, *y'know* and *I mean* operate on the other levels: *exchange*, *action*, *participation framework* and *information state*. Thus, Schiffrin (1987) points out that DMs are necessary as they contribute to the coherence of discourse through relating different components of talk in the sense that the interpretation of any component is dependent on the interpretation of the other, as she states in the following:

Since coherence is the result of integration among different components of talk, any device which simultaneously locates an utterance within several emerging contexts of discourse automatically has an integrative function. That is, if a marker acts like an instruction to consider an upcoming utterance as speaker-focused on prior text within an information state, with a simultaneous instruction to view that utterance within a particular action structure, then the result is a type of integration between those components of talk. (p. 330)

Schiffrin (1987) states that “*oh*, *well*, *and*, *but*, *or*, *so*, *because*, *now*, *then*, *I mean*, *y'know*” are discourse markers and analyses their functions in detail (p.41). For example, she states that *Oh* is a marker of information management: “it marks shifts in speaker orientation (objective and subjective) to information which occur as speakers and hearers manage the flow of information produced and received during discourse” (Schiffrin, 1987:101). *Oh* can be used for many purposes; as an exclamation or interjection, to initiate utterances, in repair initiation, in repair completion, in questions, answers and acknowledgements, as recognition receipt, as information receipt. She also

claims that although *oh* is a marker of cognitive tasks, its use may have pragmatic effects in interaction.

For example, the following illustrates *oh* as information receipt:

- Zelda: Uh::when I was a child, I lived in Glendale. It was lovely, it was beau-oh here he is! Wanna talk t'him too?
- Debby: **Oh** I'll say hello.
- Zelda: Hey, Henry, you girlfriend's here!
- Henry: [from living room] **Oh** yeh?
[enters kitchen] **Oh** how y'doin'? How are y'hh? (p.94)

The model Schiffirin (1987) proposed consists of functionally related group of items taken from other classes which can be particles (*oh, well*), conjunctions (*and, but, or, so, because*), time deictics (*now, then*), lexicalised clauses (*y'know, I mean*) and others.

Furthermore, Fraser (1999) has also studied discourse markers in detail and proposed an account of them within coherence approach similar to Schiffirin. However, some aspects suggested by Fraser are different from those of Schiffirin's. Basically, he (1999) states that discourse markers are "a class of lexical expressions drawn primarily from the syntactic classes of conjunctions, adverbs and prepositional phrases with certain exceptions" (p.831).

Fraser (1999) prefers to use 'discourse segment' as a cover term to refer to 'proposition', 'sentence', 'utterance' and 'message' and he states that discourse markers "function like a two-place relation, one argument lying in the segment they introduce, the other lying in the prior discourse: <S1. DM+S2>". However, he also mentions that this may not always be the case. For example:

He drove the truck through the parking lot and into the street. Then he almost cut me off. After that, he ran a red light. **However**, these weren't his worst offenses.

Here *however* is relevant to the segment it introduces ('These weren't his worst offenses') not with just directly prior segment (after that, he ran a red light'), but with several prior segments.

Moreover, another issue about discourse markers is their position in an utterance or a sentence. According to Fraser (1999), "almost all DMs occur in initial position (*though* being an exception), fewer occur in medial position and still fewer in final position" (p.938), as it can be seen in the following examples:

- a. Harry is old enough to drink. **However**, he can't because he has hepatitis.
- b. It is freezing outside. I will, **in spite of this**, not wear a coat.
- c. We don't have to go. I will go, **nevertheless**.

Fraser (1999) excludes some expressions from the class of discourse markers:

- a. Harry is old enough to drink, B: **Frankly**, I don't think he should.
 - b. I want a drink tonight. **Obviously**, I'm old enough.
 - c. A: We should leave fairly soon now. B: **Stupidly**, I lost the key so we can't.
- (p.942)

Within the examples above, *frankly*, *obviously* and *stupidly* do not refer a two-placed relationship between the adjacent discourse segments, but rather signal a comment, a separate message, so they are '*commentary pragmatic markers*' rather than discourse markers. Similarly, Fraser also excludes focus particles such as *even*, *only*, *just* and pause markers such as *Hum...*, *Well...*, *Oh...*, *Ahh...* as in the following examples:

- a. The exam was easy. **Even** John passed.
- b. They are fairly restrictive there. **Only** poor Republicans are allowed in.
- c. What am I going to do now? **Well...** I really don't know.
- d. A. Do you know the answer? B: **Ah ...**, I will have to think about it.

Moreover, he also excludes vocatives and interjections, which is different from Schifffrin, as they do not signal a relationship between segments, such as in the following:

- a. We shall arrive on time. B. *Sir*, I fear you are sadly mistaken.
- b. Who knows the answer. *Anyone*?
- c. The Chicago Bulls won again tonight. B. *Oh*!
- d. *Wow*! Look at that shot! (p. 943).

Fraser also discusses the grammatical status of discourse markers by stating that DMs do not form a separate syntactic category, which means that they can be expressions gathered from different classes. There are three sources of DM – *conjunction, adverbs and prepositional phrases*, of which examples can be given in the following respectively:

- a. *Since* Christmas, we have had snow every day.
- b. Bill likes to walk. *Conversely*, Sam likes to ride.
- c. You shouldn't do that. *In particular*, you shouldn't touch that brown wire.

When analyzed semantically, meaning of expressions may differ when they function as DMs. Fraser points out the necessity of discourse markers in a sentence/an utterance by stating that discourse markers do not contribute to the propositional meaning of both segments. For instance, the DMs in the following sentences may be omitted without any change in the propositional content of the segments but in that case, the hearer is left without a lexical clue for the relationship between two segments:

- a. I want to go to the movies tonight. *After all*, it's my birthday.
- b. John will try to come on time. *All the same*, he is going to be reprimanded.

This characteristic of discourse marker cannot be possible for all discourse markers. Some of the discourse markers such as *since, while, whereas* and *because* cannot be deleted because of syntactic reasons.

Moreover, according to Fraser (1999) the meaning of a DM is *procedural* not conceptual. This means that the meaning of a discourse marker is to be interpreted in relation to the prior, subject to the constraints mentioned earlier, which also means that every DM has a specific, core meaning. For example, in the following sentence, *in contrast* signals a specific contrast with the prior segment along two specific contrast areas:

John is fat. ***In contrast***, Jim is thin.

Choosing one discourse marker instead of another whose meaning is similar depends on the context. This is the main reason using *in contrast* instead of *but* or *nevertheless* as it signals a more specific contrast. The following examples can also be given to point out this feature of discourse marker:

- a. A: Harry is honest. B: **But/*In contrast**, he is not honest.
- b. He hasn't been feeling that well. **On the other hand/* In contrast**, he shouldn't have acted that way.
- c. I don't care for peas. **In contrast/*Nevertheless**, I like carrots.
- d. We started late. **Nevertheless/* In contrast**, we arrived on time. (Fraser, 1999:945)

Fraser (1999) suggests that DMs be considered as a pragmatic class as they contribute to the interpretation of an utterance rather than to its propositional content. Fraser's (1999) point of view about discourse markers is mostly similar to Schiffrin (1987) as they both discuss DMs in a coherence-based account which points out that DMs signal coherence relationships between units of talk. Moreover, they also claim that DMs are linguistic expressions constituted from several syntactic classes. However, there are also differences between Schiffrin's and Fraser's account of discourse markers.

One of the differences between Schiffrin and Fraser's account is that Schiffrin suggests that DMs contribute to 'local coherence' while Fraser claims that they contribute to 'global coherence', which means that DMs can signal the segment they introduce (S2) to any other previous segment in discourse. Another important difference is that Schiffrin analyzes discourse markers by focusing on the structural and linguistic role of DMs in contributing coherence. On the other hand, Fraser discusses the

cognitive role of DMs and suggests that DMs have procedural meaning, which is similar to Blakemore's (1987) argument that is based on relevance-based account of discourse markers.

Coherence relations are also called by some authors as rhetorical relations and a part of Rhetorical Structure Theory (RST) which is an approach to textual coherence and organization and addresses text organization by means of relations that hold between parts of a text. Taboada (2006) proposes that DMs are as signals of rhetorical relations through the study of two different corpora; a study of conversations and a study of newspaper articles and explains the coherence relations such as in the following: .

Coherence in discourse can be achieved by different means. Coherence relations-relations that hold together different parts of the discourse-are partly responsible for the perceived coherence of a text. More specifically, the recognition of coherence relations by the hearer or reader enables them to assign coherence to a text. Discourse markers guide the text receiver in the recognition of those relations. (p.567-568)

Although rhetorical relations and coherence relations have several similarities, Taboada (2006) states that "there are differences between them by pointing out that the main difference is that rhetorical relations place emphasis on the writer's intentions and the effect of the relation on the reader" (p.567). Moreover, she states that there is experimental evidence that shows that discourse markers are used in the recognition of rhetorical relations by giving examples from Haberlandt (1982) who tested reading times with marked and unmarked relations between two sentences, and found that the pairs that were marked with a discourse marker were processed faster.

Taboada (2006) clarifies how discourse markers make the relation between utterances explicit through following examples:

- (1) A. Tom quit his job.
B. He was tired of the long hours.
- (2) Tom quit his job *because* he was tired of the long hours.
- (3) Tom quit his job. He was tired of the long hours, *anyway*. (p.569)

The utterances in (1) can be interpreted as related although there is no explicit proof of it. However, the relation can be made explicit through the conjunction because as in example (2) or a different marker would void the casual relation, similarly in the case with *anyway* in (3).

As a result, it can be concluded that although there are some disagreement about specific details, generally, the coherence account of discourse markers focus more on their textual functions as they provide contextual coordinates for utterance.

2.4.2. Relevance-based Account of Discourse Markers

Another approach which is suggested by several researchers to inteprete utterances is the Relevance Theory. Diane Blakemore (1987) is the leading researcher within this approach, by pointing out the necessity of relevance relations between utterances as in the following:

Utterance interpretation is not simply a matter of identifying the proposition expressed. It is also a matter of recovering the intended contextual effects of the utterance... the coherence of a text may derive from the way in which the relevance of one segment depends on the interpretation of another. (p.134)

The Relevance Theory is based on Grice's (1975) maxims and implicatures and Sperber and Wilson's theory (2004). Paul Grice (1975) focused on the implicatures that he coined, which provide explicit account of how a certain utterance means more what is literally expressed. He also prososed four maxims such as quantity, quality, manner and relevance within the cooperative principle, all of which are used to explain the link between the utterances. Grice's (1975) central claim is that essential feature of most human communication both verbal and non-verbal is the expression and recognition of intentions. With the insights of Grice's principles, Deirdre Wilson and Dan Sperber have developed Relevance Theory. They state that the central claim of the relevance theory is that when an utterance is made, there ocur also expectations of relevance and the relevance should be precise and predictable enough to guide the listener towards the speaker's intended meaning. In order to attain the relevance, cognitive aspects are fundamental. Wilson and Sperber (2004) explain this issue of relevance of an input by stating that "in relevance-theoretic terms, other things being equal, the greater the

positive cognitive effects achieved by processing an input, the greater its relevance will be”. Therefore, there should be positive cognitive effects of an input in order to reach the relevance.

Moreover, according to the Relevance Theory, a new item of information can be taken as relevant in three ways:

1. It may allow the derivation of a contextual implication.
2. It may provide further evidence for and hence strengthen an existing assumption.
3. It may contradict an existing assumption and lead to the elimination of an assumption. (Wilson & Sperber, 2004)

In each case, the relevance depends on the contextual assumptions. Blakemore (1990) exemplifies these assumptions as in the following examples:

- (1) If David isn't here, then Barbara is in town.
- (2) David isn't here.
- (3) Barbara is in town. (p.135)

The sentences above can be interpreted in the way that the contextual assumption in (1), the information in (2) will be relevant in virtue of yielding the contextual implication in (3).

Moreover, in spoken discourse, the following two utterances are connected by the speaker by means of either intonation or the use of discourse connectives like *so*, *after all*, *moreover* or *however*.

- (a) Barbara is in town.
- (b) David isn't here.

Each connective changes the meaning of utterances. The differences can be considered in the following examples:

Barbara isn't in town. **So** David isn't here.

Barbara isn't in town. *After all*, David isn't here.

Barbara isn't in town. *Moreover*, David isn't here.

Barbara isn't in town. *However*, David isn't here. (Blakemore, 1990:136)

Jucker (1993) believes that relevance theory is “the only theory that can account for all the uses of *well* on the basis of a general theory of human communication based on cognitive principles and discusses that the notion of context is highly fundamental in relevance theory” (p.438). In order to understand how the notion ‘context’ is used in relevance theory, the following three aspects of relevance theory should be taken into consideration:

1. Every utterance comes with a guarantee of its own optimal relevance
 2. The relevant context is established as part of the utterance interpretation
 3. Discourse coherence is the outcome of negotiating relevant background.
- (Jucker, 1993:438)

In utterance interpretation, the *optimal relevance* of the particular utterance can be attained by the comprehensive context within the situation. Jucker (1993) briefly states this necessity as “the more information an individual can get out of an utterance the more relevant it will be; and the higher the processing effort needed the smaller the relevance” (p.438). Thus, the discourse coherence occurs when the hearer recognises an utterance as relevant within the context established by the immediately preceding utterance. In addition, Jucker (1993) states that cohesive devices are adequate to guarantee coherence since coherence is a function of utterance interpretation and thus, “relevance theory provides more plausible explanations for a wide range of occurrences of discourse markers and it is superior as it accounts for all the examples in the relevant literature” (p.440).

2.5. Discourse Markers and Language Teaching

Although there were several studies done on discourse markers and their functions in the past, in recent years the number of the studies (Svartvik, 1980; Östman, 1981; Schiffrin, 1986; Aijmer, 1987; Schourup, 1985; Erman, 1987) about the issue of how the use of discourse particles contribute to the pragmatic and communicative

competence of speakers or the pedagogical significance of discourse markers in language teaching has increased. Several studies have claimed that discourse markers have a fundamental place in language teaching. The studies about this particular field are mostly about the discourse markers used by native speakers of English. Although the nativeness of English is another controversial issue as it became difficult to describe the native speakers, the number of comparative studies which compare the discourse markers used by native speaker and non-native speakers is not satisfying.

The comparative studies (such as Müller; 2004; Fung and Carter, 2007) have shown that foreign language learners use fewer discourse markers when compared to native speakers and they underutilize them especially for pragmatic functions.

The necessity of the discourse markers to be learned by non-native speakers of English is salient. Lam (277) points out that discourse particles are crucial for learners to communicate successfully at the pragmatic level of interaction. Moreover, he states that:

If language learners are denied access to these critical pragmatic devices in their learning process, they may not be able to fully project their personality in the target language...learners are deprived of the right to behave and express themselves in the same way as they do in their mother tongue. The image that they could present in the second or foreign language is, at most, a partial alter ego. (p.277)

In other words, discourse markers may provide non-native learners of English gain nativeness in a foreign language in spoken or written discourse. This feeling of nativeness will help learners feel comfortable while learning another language. Through discourse markers in spoken discourse, the naturalness of talk can be attained and in written discourse, the text gains a higher level of coherence.

Hellermann and Vergun (2006) also emphasize that “many language learners have ‘grammatical’ language as the primary goal of their language learning experiences” (p. 158). However, this ‘grammatical’ target proficiency is often defined as what native speakers of the language consider accurate usage of syntax, phonology, morphology, and semantics so that the propositional content of an utterance is made clear. Thus as discourse markers are words or phrases that function within the linguistic

system to establish relationships between topics or grammatical units in discourse, the learners will have a proficiency above.

Therefore, the awareness of foreign language learners about these linguistic units should be increased. To that end, Siniajeva (2005) lists the following points to be considered while learning discourse markers for coherent speech and writing:

- how frequently they are used (*e.g. however is more frequent than nonetheless*)
- whether they generally precede clauses (*e.g. so, thus, also*) occur within clauses (*e.g. therefore*), or come at the end of the clauses (*e.g. too*)
- how they are used in relation to particular kinds of text and context (*formal, informal, written or spoken*)
- whether they can introduce or separate substantial blocks of text (*e.g. however, furthermore*), or they tend to be used with shorter stretches (*e.g. as well*). (p.6)

Increased awareness on the textual and interpersonal functions of discourse markers may be attained not necessarily through explicit teaching. Hellerman and Vergun (2007) give some pedagogical examples about how to increase the awareness such as language samples from everyday conversation between fluent speakers of the target language, giving students adequate time for pair and small-group interaction in class.

To sum up, the importance of teaching discourse markers explicitly or implicitly in language teaching has been studied by several authors. It has been acknowledged that there is a strong relationship between discourse markers and pragmatic competence. If language learners are capable of using discourse markers effectively and adequately in spoken discourse, their utterances will be much more understandable for the hearer or the listener. Therefore, the use of discourse markers provides language learners have discourse-pragmatic competence, which makes its place precise and clear in the mainstream of language teaching.

CHAPTER 3

METHODOLOGY

3.1. Introduction

This chapter includes the detailed information about the methodology of the study. Firstly, as the study is nourished from two specific corpus; the selection process of the corpus and the descriptions about them are highlighted. Then, the chapter presents data collection procedure which was necessary to gather a particular corpus of non-native speakers of English. The data collection procedure includes the explanation of the instruments, setting, participants and transcription process. Finally, the data analysis is explained in a way that shows how both corpus were analyzed to reach the objectives of the research.

3.2. Corpus Selection

As the study aims to define and mark the discourse markers that non-native speakers of English use in their spoken discourse and compare the results with the ones uttered by native speakers of English, there has been definitely a necessity for choosing two specific corpus which will constitute the basis of the research; MICASE and Research Corpus. The first corpus was a ready-made corpus, however; the second corpus is the corpus constituted for the study by the researcher which is explained in detail in the data collection section.

The first corpus that was taken as the basis and a sample for the second corpus was obtained from Michigan Corpus of Academic Spoken English (MICASE). MICASE was prepared by Rita C. Simpson, David Y.W.Lee and Sherly Leicher in 2002 and revised by Annelie Ädel in 2007. The MICASE Corpus is a spoken language corpus which is available on-line at <http://micase.umd.umich.edu/m/micase/>. The corpus consists of approximately 1.8 million words (nearly 200 hours of academic speech) including contemporary university speech within the microcosm of University of

Michigan, in Ann Arbor, Michigan, USA. It is the outcome of a research project conducted by the English Language Institute (ELI) at the University of Michigan.

The MICASE Corpus includes academic speech, but academic speech for the project means the speech which occurs in academic setting, which points out that it is not limited as not only “scholarly discussion” but also all kinds of jokes, confessions and personal anecdotes. Thus, it provides a wide range of academic speech events. Within the corpus, there are 152 speech events available; 2 Advising Sessions, 14 Colloquia (Public Lectures), 9 Discussion Sections, 4 Dissertation Defenses, 3 Interviews, 8 Lab Sections, 6 Lectures (Small and Large), 6 Meetings, 14 Office Hours, 7 Seminars, 2 Service Encounters, 11 Student Presentations, 8 Study Groups and 2 Tours.

There is an on-line search engine including all the transcripts of academic speech events recorded. For the subject of the study, “student presentations” of native speakers of English who are senior-undergraduates were chosen as they would serve more and better for the objective of the study. Student presentations were chosen as the academic speech that can be transcribed more effectively and properly and analyzed as the sample of spoken discourse both of native and non-native speakers of English. Moreover, the permission for using these transcripts of MICASE was taken from its authors.

3.3. Data Collection

The research corpus has been constituted by the researcher. After the process of searching for the appropriate corpus that can be used as comparable with the research corpus, it has been aimed that a similar corpus should be organized by the researcher. In order to search for the objectives of the study, there should be a corpus of non-native speakers of English. The corpus which was organized for the research is presented in detail in the following section.

3.3.1. Participants

The participants in the study are four-grade undergraduate students studying in Department of English Language Teaching at Gazi University, Turkey. The students studying in Department of English Language Teaching are educated to be English teachers through the courses specialised with the field. In particular, the number of the students participated in the research are twenty.

3.3.2. Setting

The data were collected in two different courses of the four grade curriculum of 2010-2011 Academic Year Spring Semester, which is “Sociolinguistics and Language Teaching” and “Pragmatics and Language Teaching” in Department of English Language Teaching of Gazi University, Turkey which is one of the leading and populous departments within the field in Turkey with more than 1250 bachelors and nearly 40 graduate students. The course of *Sociolinguistics and Language Teaching* includes how the use of language differ in different social domains, the appropriacy of language according to certain social context as well as the correctness of the grammar and thus proposing methods to use these issues in language teaching. Moreover, the objectives of the course, *Pragmatics and Language Teaching*, are to gain a perspective of language which will lead trainees to display a better competency in handling and teaching English; to cove a wide range of pragmatics studies so that trainees can draw an effective picture of what language is and why contemporary language teaching approaches and principles counts.

3.3.3. Instruments and Procedures for Data Collection

As conducting an analysis in spoken discourse requires different methods and instruments than written discourse, the following procedures were followed to compose the research corpus.

3.3.3.1. Recording

For the purpose of collecting the data within spoken discourse of non-native speakers of English, the student presentations of Turkish undergraduate students were first audio-recorded. Since the student presentations were conducted in a classroom which is an indoor setting and as the number of students within the classrooms were higher, the type of recording equipment was selected accordingly. A portable audio-recorder of professional quality was used for recording the student presentations. Furthermore, the device was tested before the recording sessions and then used to record the participants' presentations. The participants and the instructor of the course were asked for their permission and the recordings were conducted with the consent of the participants. During recording sessions, an observation sheet (Appendix A) was used to get necessary information about the recording such as the title of the presentation, date, gender of the presenter, grade of the presenter, duration of the presentation and a section for possible notes that could be relevant about the presentation.

At the end of each recording session, the quality of the recording was confirmed in order to determine whether the recordings were intelligible or not, which was of utmost importance for the correctness while transcribing. After each recording session, the recordings were transferred to the computer and saved as sound files. This procedure of digital copies of recordings was done in order to facilitate the transcription and to keep the recordings for backup.

The total number of recorded student presentations is 30. The number of the recordings were higher than the recordings to be transcribed for the analysis. The underlying reason is that there could be any kind of recording problem, which might cause that the research could not reach an optimal number of presentations. Another reason was that there would be a selection process among the recorded student presentations.

3.3.3.2. Transcription

As is done in spoken corpus studies, after the recording procedure, the transcription session starts. In corpus studies, the transcription process should be carried out carefully. Du Bois (1991) defines discourse transcription as “the process of creating a representation in writing of a speech event so as to make it accessible to discourse research” (p.72). Thus, the transcription of spoken discourse makes it accessible to conduct a study on it.

Moreover, “transcription documents language use, but language use is attested equally in written discourse, which has the advantage of being easy to obtain without transcribing” (Du Bois, 1991, p.73). Thus, conducting research on spoken discourse requires much more effort. However, spoken discourse enables the process of the production of language to be more accessible to observer. For example, “hesitations, pauses, glottal constrictions, false starts and numerous subtle evidences are observable in speech and provide clues to how participants mobilize resources to plan and produce their utterances, and to how they negotiate with each other the ongoing social interaction.” In order to make spoken discourse accessible for a research, transcription should be used. Thus, Du Bois (1991) states that “a transcription of spoken discourse can provide a broad array of information about these and other aspects of language, with powerful implications for grammar, semantics, pragmatics, cognition, social interaction, culture, and other domains that meet at the crossroads of discourse” (p.73).

In order to carry out an effective transcription process, there should be particular transcription conventions to be set on. As the research corpus is constituted by taking MICASE Corpus as the basis for the study, the transcription process was conducted in a similar way and manner with MICASE Corpus. The transcription of the research corpus was done according to MICASE orthographic transcription conventions and mark-up system (see Appendix B) which are organized to allow for ease of readability, while including enough detail to ensure adequate comprehension from the text of the transcript alone.

After setting the conventions for the transcription, the transcribing process which was so arduous and time-consuming started. The selected recordings were transcribed by using the conventions stated above and directly into a computer file using a computer program that was originally developed for the MICASE Project, called Sound Scriber. This program for Windows which has various user-configurable features was created by Erick Breck in 1998 at the University of Michigan and its primary function is to carry out the transcription of digitized sound files easily and effectively. The program is available for free and includes features specifically for transcription such as keystrokes to control the program while working in another window (e.g. word processor, SGML editor, etc.), variable speed playback, and a feature called "walking", in addition to normal playback features. As it is stated in its website, by the help of this particular feature "walking" which plays a small stretch of the file several times and then advances to a new piece, overlapping slightly with the previous one, it is possible to transcribe continuously without having to manually pause or rewind the recording.

Another important point during transcription is the fact there were some speech errors made by the participants. They were not corrected and were transcribed as how they had actually occurred.

3.4. Corpus Description

As mentioned above, the research was conducted on two corpus; MICASE and research corpus. Within MICASE, a sub-corpora has been chosen through "Browse Transcripts" section. The speaker attributes for this particular sub-corpora are "senior undergraduate" as academic position and role, "American English native speaker" as native speaker status. As for transcript attributes, "student presentations" has been chosen as speech event type and the option of all was chosen as academic division to have a comparable number of transcripts within the data analysis.

Although there are six transcripts in sub-corpora, it does not mean that there are only six presentations. In each transcript, there are several students as speakers present a specific topic consecutively. Moreover, after all transcripts have been examined and their on-line audio files have been listened, four transcripts have been chosen according to their disciplines. The presentations done by the students in the fields of social

sciences and education and humanities were chosen in order to have comparable context in transcripts with the research corpus. In the following table, the relevant information about the transcripts taken from MICASE are given in detail. Primary discourse modes of the transcripts are also presented by ranging from mostly monologic, mostly interactive and mixed interactivity rating. Moreover, total duration is **300** minutes with total word counted as **41.173**.

Table 3: Description of MICASE transcripts

No	Title of the presentation	Academic Division	Primary Discourse Mode	Duration (in minutes)	Total word count
1	Second Language Acquisition	Humanities	Mixed	69	7.384
2	Bilingualism	Humanities	Mostly Interactive	99	14.572
3	Multicultural Issues in Education	Social Sciences and Education	Mostly Monologic	66	10.195
4	Black Media	Social Sciences and Education	Mostly Interactive	66	9.022
			TOTAL	300	41.173

In order to develop the small-scale research corpus of non-native speakers of English, thirty presentations were recorded. All of the presentations were listened to and a selected fragment of totally 315 minutes by twenty speakers was taken as the main focus for analysis and interpretation as it constitutes the most representative and richest section in terms of oral interaction among the participants. As the table below presents, the primary discourse modes are mostly mixed and mostly monologic and total duration is **315** minutes with total words counted as **34.420**.

Table 4: Description of transcripts of the research corpus

No	Name of the presentation	Primary Discourse Mode	Duration (in minutes)	Total word count
1	Sociolinguistics:social class	Mostly Monologic	15	1.354
2	Social class:vowels	Mixed	11	1.294
3	Gender and age	Mostly Monologic	11	1.279
4	Pragmatics: Implicatures	Mostly Monologic	13	1.398
5	Pragmatics and Indirectness	Monologic	19	2.347

6	Indirectness	Mixed	21	2.358
7	Maxims	Mixed	15	1.664
8	Pragmatics and Indirectness	Mixed	10	1.138
9	Pragmatics and Power	Monologic	24	2.054
10	The reasons of indirectness	Mixed	8	838
11	Language and age	Mostly Monologic	16	1.501
12	Conflicts and register	Mixed	23	2.816
13	Size of imposition	Mixed	11	1.369
14	Accommodation theory	Mixed	33	4.001
15	Sociolinguistics:colloquial style	Mixed	19	1.884
16	Sociolinguistics:register	Mostly Monologic	16	1.458
17	Pragmatics and Indirectness	Mixed	15	1.406
18	Pragmatics: The construction of meaning	Mostly Monologic	10	1.452
19	Pragmatics: Speech Acts	Mostly Monologic	17	1.768
20	The construction of meaning	Monologic	8	1.041
TOTAL			315	34.420

3.5. Data Analysis

The research is based on mixed method data analysis; quantitative and qualitative analysis. The results obtained from the transcripts were analysed mostly quantitatively. Quantitative analysis was conducted by the use of descriptive statistics. Descriptive statistics were used to display the occurrences and distribution of discourse markers in the discourse through lexical size and frequency counts.

There were several procedural steps performed within the quantitative analysis of the research, which lead the research to display the results. First of all, all transcripts were analyzed in detail regarding with which words or phrases were qualified as discourse markers. During this part of analysis, the functions of discourse markers proposed by Schiffrin (1987), Brinton (1996), Müller (2005) and Fung and Carter (2007) were taken as the basis to search for discourse markers. Thus, the list of discourse markers was composed. Meanwhile, the analysis was also examined by an expert who has a doctoral degree on discourse analysis.

Then, in order to display and process results of a corpus of language in an effective way and reach to the analysis for descriptive statistics, a concordancer program was used. As Sun & Wang (2003) describe, “concordancers have been shown

to be an effective aid in the acquisition of a second or foreign language, facilitating the learning of vocabulary, collocations, grammar and writing styles” (p.6).

The program used for the research is called *AntConc* which is a freeware multipurpose corpus analysis toolkit designed by Laurence Anthony at Waseda University for specific use in the classroom. As Laurence Anthony states that AntConc includes a powerful concordancer, word and keyword frequency generators, tools for cluster and lexical bundle analysis, and a word distribution plot. Thus, with the help of AntConc, each discourse marker was displayed within the concordance lines through which each one of the instances in which discourse markers occur was analyzed. As AntConc lists all the instances of that specific discourse marker, there were some words or phrases listed within the concordance lines which may serve other functions different from their use as discourse markers. In these kind of instances, an analysis for differentiating discourse markers and non-discourse markers was performed. In the following step, each item which is identified as discourse markers is counted within each transcript respectively, in both research corpus and MICASE corpus. One of the fundamental points to be mentioned for data analysis is that the significance of this research is that all the items that serve the function of a discourse marker are identified. In other words, not only a limited of discourse markers identified beforehand were analyzed within the corpus. Afterall, a comparative analysis was done for the purpose of displaying the contrastive frequency, which lead to discussion upon the results. Therefore, the quantitative side of the analysis was finalised.

Another part of the data analysis is qualitative analysis which includes the study of core functional paradigm of discourse markers in academic discourse based on the multi-categorical model proposed by Fung and Carter (2007). The categories proposed are interpersonal, referential, structural and cognitive category. Thus, while analyzing qualitatively the results of the research, examples are presented from the research corpus and MICASE to display how the discourse markers uttered serve for which function.

All in all, the data analysis of the research was conducted both quantitatively and qualitatively. Both types of analysis, the results of which is presented in the following chapter, are necessary to carry out a research on discourse analysis. The nature of the

discourse analysis requires qualitative aspects as discourse includes several functions of certain items to be analyzed. Thus, it is not possible to present the results of the analysis in spoken discourse through only numeric tables of frequencies. On the other hand, there is also a need for quantitative results of the data to display concrete outcomes of the discourse items and to compare the results of native speakers and non-native speakers.

CHAPTER 4

ANALYSIS OF RESULTS

4.1. Introduction

In this chapter, the results of the analysis of discourse markers in the research corpus and MICASE are presented in order to identify the similarities or differences of the discourse markers and the reasons why these may occur. All these results are to be used during the interpretation process of the analysis. For these purposes, numeric tables including the items to be analyzed were developed. It has been aimed that these tables would definitely shed light on the overall analysis and allow the researcher to conclude with specific generalisations about the use of discourse markers in American and Turkish spoken discourse in pedagogic settings.

This chapter first presents how the process to differentiate the items that function as discourse markers and the ones do not function as discourse markers. Then, in the next section, the results obtained from the analysis in the research corpus and MICASE corpus are displayed and explained one by one through tables in succession. After each tables for both corpus are shown, to highlight the similarities and differences between both corpus, a comparative quantitative analysis is presented through frequency analysis. Then, to support the research and make the analysis richer, a qualitative analysis is conducted through extracts taken from both corpus with particular functions. Finally, the chapter ends with the discussion of the results obtained by highlighting particular comparative comments which contribute significantly to the pedagogical implications of the research.

4.2. Identifying Discourse Markers and Non-discourse Markers in Research Corpus and MICASE Corpus

In order to carry out the analysis of the results obtained from the transcripts, one of fundamental steps is to form criteria for identifying what kind of utterances can be named as discourse markers. As it is presented, the research is based on the functions of discourse markers proposed by Schiffrin (1987), Brinton (1996), Müller (2005) and

Fung and Carter (2007), as mentioned in the review of literature. Thus, during the process identifying the discourse markers in both corpus, some instances of the item in question were excluded from the research. Moreover, the items uttered by the speaker while reading the examples about the topic or examples taken from dialogues were also excluded since they are not uttered originally by the speaker and s/he just reads the text. The following extracts are given as examples from the research corpus to display how each item is differentiated.

Well is one of those items which should be treated carefully during analysis. *Well* is itemized when it is used to fulfill the function of denoting thinking process, as a hesitation marker or used to open and close the topics. However, in cases when *well* collocates with “very” as an adverb qualifying an adjective, it cannot be identified as discourse marker as it does not fulfill any discourse marker function. For example, in the following extract *well* is used as an adverb so these kind of instances are excluded.

...In other words, if you close uhh if you feel close to someone because that person is related to you, or you know him or her **well** uhh or he or she is similar to you in terms of your age, social class, occupation etc. you feel -ness uhh you feel less need to employ indirectness... (Transcript 6, Research Corpus)

However, in such cases in the following extracts *well* functions as a discourse marker to turn to another issue and these instances are included to the analysis.

..., he simplifies her speech her uhh speed of speech, pronunciations, so he converges her speech towards the less linguistic proficiency of her students. **well** uhh please open your books, turn uhh turn to page at two hundred thirty one. (Transcript 12, Research Corpus)

And and *or* are the other discourse markers to be itemized selectively. *And* is not considered as a discourse marker when it is used to join two words together, to enumerate items as it occurs in the sentence structure noun + noun or when showing one thing happens after another, as in the following extracts.

..., different people prefer different uhh styles **and** registers. uhh more generally, register is also used to indicate degrees of formality in language use. (Transcript 12, Research Corpus)

...when you look at the next example uhh David tea or coffee and Jenny says "yes please" **and** then they would say coffee Jenny "thank you". here we see that because Jenny says "yes please". (Transcript 20, Research Corpus)

In the fourth extract, **and** collocates with **then** and is used as discourse marker to continue the topic. Moreover when **and** is used to shift the topic, to display contrast or as an additive marker before a short or long pause to add something new to the ongoing topic by joining two clauses together. So, in these instances, the item meets the criteria to be a discourse marker.

A similar process is conducted for **or**. For example, in the following extracts, the instances for **or** as a discourse marker and as a non-discourse marker can be found. Although, **or** in (a) is used just between two personal pronouns to display the alternatives and does not fulfill any kind of functions of a discourse marker, or in (b) is a conjunction to display two different clauses.

..., if you close uhh if you feel close to someone because that person is related to you, or you know him **or (a)** her well uhh **or (b)** he or she is similar to you in terms of your age, social class, ... (Transcript 6, Research Corpus)

uhh another type is legitim- legitimate power, this is uhh because of role uhh age **or** status a person has and uhh a person has the right prescribe or request think about your mothers. (Transcript 9, Research Corpus)

Moreover, another discourse marker analyzed in detail is **so**. **So** functions as a discourse marker when it is used as an opening frame marker, to summarize opinions, to shift the topic, to continue the topic or to display a consequence. For example, in the following extract **so** is used to shift the topic.

...these are the focus of interest are often heavenly modified all after the noun as in A and before the noun as in B and B. do you see? yes, **so** let's have a look at our routines and formulas.

(Transcript 22, Research Corpus)

However, in the following extract, **so** is used to qualify the adjective *upset* and functions as an adverb of degree and manner. Therefore, these kinds of instances are also excluded.

... look at the example. little girl looks **so** upset while looking at a mirror. her father realises that there is a problem...

(Transcript 5, Research Corpus)

Like is another discourse marker which is used in spoken discourse as an exemplifier, (as in the following extract) to elaborate the existing propositional meaning to make the intention clearer or to mark an approximate marker.

... you see the more frequently they use the multiple negation more teenagers. and teenagers use more multiple negation than adult. okay. **like** slang vernacular forms act as solidarity markers, they can indicate membership of closely social groups.

(Transcript 11, Research Corpus)

However, the item **like** also is a verb used to show the attitude of one person in a positive way, in the following extract . So these instances are also excluded.

...ask a child i- if he would **like** a pie. actually, the answer is exactly yes... (Transcript 11, Research Corpus)

Furthermore, **right** is another common discourse marker used to signal a discourse boundary where a topic ends and another begins or to make sure whether the utterances is clear and understood, as in the following extract. On the other hand, **right** as nouns such as “right to punish, right to say” can not be itemized as discourse markers.

...ask a child i- if he would like a pie. actually, the answer is exactly yes. it's quite clear. **right?** so, the question is unnecessary. [SS: hmm.]

Therefore, the examples given above present clearly how each items in concordance lines is analyzed and differentiated whether it functions as a discourse marker or not. In particular, this process has been conducted for each instances for each item.

4.3. Findings of Analysis in the Research Corpus

As the research aims at providing comparative analysis about the discourse markers used by non-native speakers of English and native speakers of English, the first step for analysis should be to present the results of analysis conducted through the research corpus. After transcription process of the student presentations of Turkish non-native speakers of English, the tokens of discourse markers were identified one by one in each transcript, and this is followed by presenting the tokens identified in numeric tables. Then, total word count was estimated in each transcript to be used in frequency analysis. In this process, it was noted to include only the words that the speakers utter during their presentation in word count. All the other items in the transcripts were excluded. The word counts of each transcript were used to display the frequency of the items identified in transcripts. In other words, the results were presented in frequency analysis of descriptive statistics by using the variables of the number of occurrences of each lexical items, namely, discourse markers and the number of word count of each transcript. (See table below). Discourse markers are displayed in the tables for every 100 words in the corpus. This process was conducted for twenty transcripts of Turkish non-native speakers' presentations. After all transcripts were analyzed and displayed in numeric tables, a table including all the items identified in twenty transcripts was organized to present the total results of the frequency analysis.

As it is not possible to display the results of all transcripts in one table, there are two separate tables below in which all discourse markers used in each transcript are displayed.

The table below presents the number of occurrences of each discourse marker in the first ten transcripts.

Table 5: Occurrences of Discourse Markers in the first ten transcripts in the research corpus

Word Count :		1.354	1.294	1.279	1.398	2.347	2.358	1.664	1.138	2.054	838
Transcript No :		1	2	3	4	5	6	7	8	9	10
DM		Occ	Occ	Occ	Occ	Occ	Occ	Occ	Occ	Occ	Occ
1	uhh	56	37	20	72	77	106	48	44	52	25
2	okay	10	4	2	6	0	5	30	7	5	1
3	yes	8	12	8	7	0	16	13	14	6	6
4	firstly, first, first of all	0	1	1	0	0	2	2	0	2	1
5	and	14	23	30	45	33	26	19	17	49	29
6	so	1	4	8	7	14	15	18	8	15	5
7	but	6	2	8	9	6	11	14	5	13	5
8	let's...	0	0	4	0	4	9	10	9	5	3
9	for example	7	4	12	5	8	14	3	4	0	0
10	or	4	1	2	3	14	12	6	5	6	1
11	like	2	1	5	8	3	1	4	1	0	0
12	also	2	2	3	0	11	2	5	4	4	4
13	you see, see	0	0	1	0	0	2	0	0	0	1
14	however	0	0	1	2	3	2	0	0	0	0
15	yet	0	0	1	0	0	0	0	0	0	0
16	generally	0	0	2	0	0	0	0	0	0	0
17	as	5	6	1	0	10	0	2	2	3	6
18	even	0	0	1	0	4	0	0	1	0	1
19	such as	1	1	1	0	0	1	0	1	0	0
20	hih, hi-huh	0	1	2	0	0	0	4	0	0	0
21	only	1	0	3	3	0	0	0	0	0	0
22	then	0	2	0	3	3	1	0	1	7	0
23	you know	0	0	0	0	5	0	0	0	0	0
24	specifically	0	0	0	0	2	0	0	0	0	0
25	as you see	3	6	0	2	4	1	0	0	5	1
26	as an example	0	0	0	0	2	1	0	0	0	0
27	as a brief	0	0	0	0	6	0	0	0	0	0
28	hmm	5	0	0	0	0	0	7	3	1	2
29	yep	1	0	0	0	0	0	0	0	0	0
30	umm	7	6	1	0	0	0	10	12	0	0
31	as i mean	1	0	0	0	0	0	0	0	0	0
32	because	2	0	0	0	5	14	6	5	18	0
33	by the way	1	0	0	0	0	0	0	0	0	0
34	still	0	1	0	0	0	0	0	0	0	0
35	i mean	0	4	0	0	0	3	0	0	0	0
36	although	0	1	0	0	0	0	0	0	0	0
37	actually	0	1	0	0	0	0	1	0	0	0
38	clearly	0	0	0	1	1	4	0	1	0	0
39	now	0	0	0	5	0	3	1	0	4	0

40	obviously	0	0	0	2	0	1	2	0	0	0
41	in fact	0	0	0	1	0	0	0	0	0	0
42	this is to say	0	0	0	1	1	0	0	0	0	0
43	as a result of	0	0	0	0	2	0	0	0	0	0
44	simply	0	0	0	0	2	0	0	0	0	0
45	really	0	0	0	0	0	1	2	1	4	0
46	probably	0	0	0	0	0	1	0	0	0	0
47	as you know	0	0	0	0	0	1	0	0	0	0
48	in other words	0	0	0	0	0	1	0	0	0	0
49	i think	0	0	0	0	0	1	2	0	0	0
50	just	0	0	0	0	0	1	4	1	0	0
51	sort of	0	0	0	0	0	1	0	0	0	0
52	right	0	0	0	0	0	0	7	0	0	0
53	anyway	0	0	0	0	0	0	1	0	0	0
54	well	0	0	0	0	0	0	1	1	0	0
55	exactly	0	0	0	0	0	0	5	0	0	0
56	indeed	0	0	0	0	0	0	1	1	0	0
57	yeah	0	0	0	0	0	0	0	9	0	1
58	moreover	0	0	0	0	0	0	0	1	0	0
59	on the other hand	0	0	0	0	0	0	0	1	0	0
60	maybe	0	0	0	0	0	0	0	3	0	0
61	otherwise	0	0	0	0	0	0	0	1	0	0
62	lastly	0	0	0	0	0	0	0	1	0	0
63	of course	0	0	0	0	0	0	0	2	2	0

The table below presents the number of occurrences of each discourse marker in the second ten transcripts.

Table 6: Occurrences of discourse markers in the second ten transcripts in the research corpus

Word Count :		1.501	2.816	1.369	4.001	1.884	1.458	1.406	1.452	1.768	1.041
Transcript No :		11	12	13	14	15	16	17	18	19	20
DM		Occ	Occ	Occ	Occ	Occ	Occ	Occ	Occ	Occ	Occ
1	yes	14	8	13	21	8	9	6	3	4	4
2	uhh	34	216	41	132	59	38	115	49	143	37
3	like	8	8	0	0	0	0	2	0	4	0
4	or	3	7	4	4	1	0	4	6	3	3
5	as	5	11	2	8	10	21	0	0	2	2
6	but	1	13	7	13	10	2	9	9	8	2
7	first, first of all	2	1	0	8	1	0	0	0	0	0
8	then	3	4	1	3	2	1	0	0	2	4
9	also	4	2	0	8	3	2	7	2	0	0
10	now	3	1	2	4	0	0	0	0	2	0
11	and	31	59	18	54	35	41	39	33	25	17
12	for example	2	17	9	8	2	2	2	3	8	4
13	such as	3	0	0	0	3	2	0	0	0	0

These two tables are gathered in a way that the occurrences of common discourse markers are summed up and different discourse markers are displayed again within the following table.

Table 7: Occurences and frequencies of the discourse markers in the research corpus

Total Word Count : 34.420			
No	Discourse Marker	Occurence	Percent
1	uhh	1.401	4,07%
2	and	637	1,85%
3	so	186	0,54%
4	yes	180	0,52%
5	but	153	0,44%
6	umm	119	0,35%
7	for example	114	0,33%
8	let's...	63	0,18%
9	as	96	0,28%
10	because	91	0,26%
11	or	89	0,26%
12	okay	99	0,29%
13	also	65	0,19%
14	like	47	0,14%
15	then	37	0,11%
16	hmm	34	0,10%
17	as you see	32	0,09%
18	hih, hi-huh	27	0,08%
19	just	26	0,08%
20	now	25	0,07%
21	yeah	25	0,07%
22	however	23	0,07%
23	firstly, first, first of all	21	0,06%
24	really	19	0,06%
25	actually	15	0,04%
26	of course	14	0,04%
27	such as	13	0,04%
28	you know	13	0,04%
29	clearly	11	0,03%
30	i mean	10	0,03%
31	you see, see	10	0,03%
32	alright	8	0,02%
33	by the way	7	0,02%
34	even	7	0,02%
35	exactly	7	0,02%
36	for this/that reason; for these reasons	7	0,02%

37	i think	7	0,02%
38	only	7	0,02%
39	right	7	0,02%
40	as a brief	6	0,02%
41	generally	6	0,02%
42	i guess	5	0,01%
43	obviously	5	0,01%
44	probably	5	0,01%
45	in other words	4	0,01%
46	well	4	0,01%
47	as an example	3	0,01%
48	especially	3	0,01%
49	maybe	3	0,01%
50	moreover	3	0,01%
51	although	2	0,01%
52	as a result of	2	0,01%
53	in fact	2	0,01%
54	in this way	2	0,01%
55	indeed	2	0,01%
56	just like	2	0,01%
57	lastly	2	0,01%
58	otherwise	2	0,01%
59	simply	2	0,01%
60	specifically	2	0,01%
61	this is to say	2	0,01%
62	above all	1	0,00%
63	absolutely	1	0,00%
64	anyway	1	0,00%
65	as i mean	1	0,00%
66	as you know	1	0,00%
67	even though	1	0,00%
68	eventually	1	0,00%
69	finally	1	0,00%
70	for instance	1	0,00%
71	in short	1	0,00%
72	kind of	1	0,00%
73	on the other hand	1	0,00%
74	so far	1	0,00%
75	sort of	1	0,00%
76	still	1	0,00%
77	to sum up	1	0,00%
78	yep	1	0,00%
79	yet	1	0,00%
TOTAL :		3.839	11,15%

As the table above shows, total word count of twenty transcripts are **34.420**, which was taken as the basis for calculating the frequency of each discourse marker.

Within 34.420 words uttered by non-native speakers, seventy nine different discourse markers were determined. However, the number of occurrences of each discourse marker differ particularly. Moreover, total number of occurrences of all discourse markers are **3.839**, which compose **11,15 %** of the total words.

The table clearly displays that the most frequent item identified within the research corpus is **uhh**, which significantly composes the majority of the discourse markers. The item **uhh** has 4,07 % out of a total of 11,15 % and it was used 1401 times by Turkish non-native speakers of English. Furthermore, the discourse marker **and** is at the second rank of the list and it composes a larger number of discourse markers with 637 occurrences (1,85 %) within the research corpus. The discourse marker **and** is followed by the item **so** with 186 hits which constitutes 0,54 %; **yes** with 180 hits by having 0,52 % and **but** with 153 hits (0,44 %).

Additionally, the other first five discourse markers are **umm** (0,35 %), **for example** (0,33 %), **let's** (with which starts, such as **let's start, let's do...**) with 0,18 %), **as** (0,28 %) and **because** (0,26 %).

Through the list, it is noticeable that there are several discourse markers used with **as** such as **as you see, as a brief, as i mean, as you know, as a result of** and **as an example**. **As** was counted separately from these markers and these phrases were identified as discourse markers and their occurrences were counted individually. Among these items, **as you see** has the largest number of occurrences with 32 hits by having 0,09 %.

Moreover, within the list, some markers are displayed together; for example, the group of **huh, hi-huh, uh-huh**; the group of **firstly, first, first of all** and the groups of **for this/that reason, for these reasons** are the same discourse markers with slight changes in their word forms.

Another point to be taken into account is that after the first 17 items in the list, the frequencies of the discourse markers display similarities group by group. In particular, the frequencies of the group **huh, hi-huh, uh-huh** and **just** are 0,08 %; the frequencies of **now, yeah** and **however** are 0,07 % , the frequencies of the group **firstly,**

first, first of all and **really** are 0,06 %, the frequencies of *actually, of course, such as* and *you know* are 0,04 % and the frequencies of *clearly, i mean* and *you see, see* are 0,03 % although each discourse marker have slightly different occurrences.

As the list of discourse markers present, there are 76 different discourse markers used in the research corpus. However, it is clearly seen that 48 discourse markers are under the percentage of 0,03 %. The occurrences of these discourse markers are less frequent and range from 8 and 1 hits. In particular, the discourse markers representing 0,02 % each are *alright, by the way, even, exactly, for this/that reason* (including *for these reasons*), *i think, only, right, as a brief, generally, i guess*. Moreover, the discourse markers in the following; *i guess, obviously, probably, in other words, well, as an example, especially, maybe, moreover, although, as a result of, in fact, in this way, indeed, just like, lastly, otherwise, simply, specifically* and *this is to say* represent only 0,01 % each in the research corpus.

Furthermore, among the discourse markers in the list, there are also several discourse markers that have only one hit in the research corpus. These are *above all, absolutely, anyway, i mean, as you know, even though, eventually, finally, for instance, in short, kind of, on the other hand, so far, sort of, still, to sum up, yep* and *yet*. Since the occurrences of these items are just one, they are not displayed within the frequency rate. In other words, they don't represent any particular contribution to the frequency analysis.

As the brief summary of the results above shows, to give a clear picture of the discourse markers used by Turkish non-native speakers of English, the graph below also presents the frequency rates of top 20 most frequent discourse markers within the total number of occurrences of discourse markers, which is **3.839**.

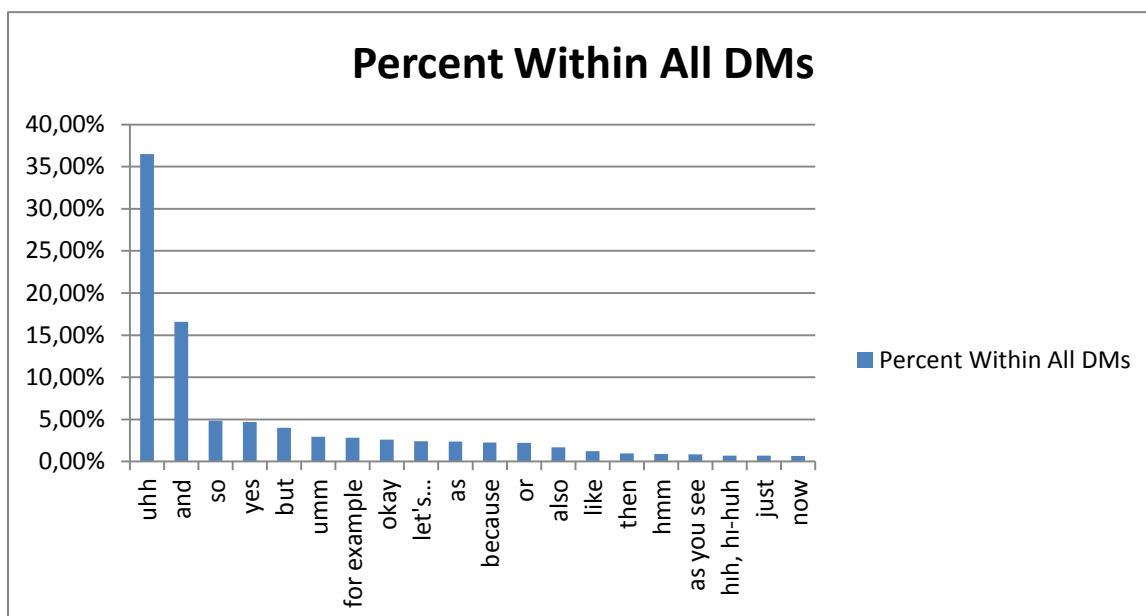


Figure 2: First 20 most frequent discourse markers in the research corpus

4.4. Results of Analysis in MICASE Corpus

The results of the analysis in MICASE Corpus were obtained by following a similar process with the one conducted in displaying the results of the analysis in the research corpus in order to arrange tools for a comparative analysis. As the transcripts to be analyzed were taken directly from MICASE corpus, the tokens of discourse markers were identified again one by one in each transcript, and presented numeric tables. Similarly, total word count was estimated in each transcript to be used in frequency analysis by considering only the words that the speakers utter during their presentation. The word counts of each transcript were used to display the frequency of the items identified in transcripts. (See table below). This process was conducted for four transcripts of the presentations of native speakers of English.

The table below presents the list of the discourse markers identified in the corpus of native speakers and their occurrences within each corpus.

Table 8: Occurrences of discourse markers in MICASE transcripts

		Word Count :	7.384	14.572	10.195	9.022
		Transcript No :	1	2	3	4
		DM :	Occ	Occ	Occ	Occ
1	so		80	137	80	62
2	just		50	73	80	63
3	oh		14	16	10	10
4	but		60	110	57	68
5	yeah		28	38	11	21
6	alright		1	2	3	3
7	okay		37	57	7	52
8	first, first one		24	3	6	3
9	sure		8	1	1	0
10	you know		14	60	42	38
11	umm		151	366	266	106
12	well		17	31	0	20
13	and		169	332	372	207
14	such as		3	1	0	1
15	uhh		38	84	96	26
16	now		6	14	11	24
17	i think		16	37	17	19
18	maybe		15	31	12	4
19	because		22	61	26	22
20	for example		5	3	0	0
21	anyway		2	1	4	2
22	quite		2	1	0	0
23	actually		12	53	11	10
24	as		18	33	14	14
25	really		19	50	50	16
26	frequently		3	0	0	0
27	then		25	30	53	21
28	also		2	48	6	13
29	let's...		6	12	3	15
30	cuz		5	9	24	15
31	second one..		10	0	0	0
32	kind of		16	24	35	6
33	later		2	6	2	4
34	you see		1	6	1	0
35	even though		6	2	0	1
36	or		18	37	22	18
37	i guess		2	14	10	1
38	i mean		7	21	18	16
39	mhm, uhuh		2	49	0	12
40	definitely		4	8	8	3

41	even	13	12	12	16
42	in conclusion	1	0	1	1
43	extremely	2	1	0	1
44	however	2	1	0	0
45	needless to say	1	0	0	0
46	generally	1	0	0	0
47	like	26	92	157	185
48	next one	2	0	0	1
49	probably	4	7	5	1
50	pretty	5	11	15	3
51	sort of/sorta	3	11	4	1
52	as you can see	1	0	0	0
53	basically	6	41	8	9
54	great	6	0	3	0
55	briefly	1	1	0	0
56	especially	3	7	0	1
57	only	10	26	6	7
58	instead of	5	2	0	0
59	i believe	1	2	0	2
60	partially	1	0	0	0
61	although	4	4	0	1
62	initially	1	1	0	1
63	previously	1	0	0	0
64	fairly	1	1	0	0
65	in general	3	0	0	0
66	kinda	1	2	13	7
67	as i said before	1	0	0	0
68	i know	2	2	7	6
69	yet	2	4	0	0
70	though	9	9	1	5
71	indeed	0	1	0	0
72	obviously	0	2	1	0
73	whereas	0	3	0	1
74	so far	0	4	1	0
75	in that sense	0	4	0	0
76	right	4	38	6	15
77	totally	0	3	3	2
78	otherwise	0	1	0	0
79	yes	0	34	0	1
80	mainly	0	5	2	0
81	as a result	0	2	0	0
82	finally	0	1	1	1
83	stuff	0	8	17	4
84	exactly	1	10	2	2
85	recently	0	1	0	2
86	specially	0	2	0	0

87	for instance	0	8	0	0
88	in fact	0	3	3	0
89	just like	1	2	4	3
90	specifically	0	1	1	0
91	afterwards	0	0	2	0
92	as well	0	0	0	5
93	whatsoever	0	0	0	1
94	eventually	0	0	0	2
95	still	0	0	0	26
96	at the same time	0	0	0	5
97	lemme	0	0	0	1
98	of course	0	0	0	4
99	right now	0	0	0	3
100	somehow	0	0	0	0
101	at the end	0	0	0	0
102	considerably	0	0	0	0
103	essentially	0	0	0	0
104	simply	0	0	0	0
105	in other words	0	0	0	0
106	by the way	0	0	0	0

The results obtained through the table above were gathered in one table to shed light on particularly which items are more frequent or less frequent, as in the following table.

Table 9: Occurrences and frequencies of the discourse markers in MICASE corpus

		Total Word Count : 41.173	
No	Discourse Marker	Occurrence	Percent
1	and	1.519	3,69%
2	umm	1.232	2,99%
3	like	612	1,49%
4	uhh	591	1,44%
5	so	541	1,31%
6	but	383	0,93%
7	just	367	0,89%
8	then	216	0,52%
9	okay	211	0,51%
10	you know	205	0,50%
11	really	188	0,46%
12	because	179	0,43%
13	yeah	147	0,36%
14	or	131	0,32%
15	also	126	0,31%

16	actually	116	0,28%
17	i mean	116	0,28%
18	as	115	0,28%
19	right	110	0,27%
20	i think	101	0,25%
21	kind of/kinda	122	0,24%
22	basically	94	0,23%
23	now	82	0,20%
24	well	81	0,20%
25	cuz	73	0,18%
26	even	69	0,17%
27	only	69	0,17%
28	maybe	68	0,17%
29	oh	67	0,16%
30	mhm, uhuh	64	0,16%
31	alright	53	0,13%
32	pretty	49	0,12%
33	stuff	49	0,12%
34	still	48	0,12%
35	let's...	43	0,10%
36	first, first one	40	0,10%
37	yes	39	0,09%
38	i guess	31	0,08%
39	probably	30	0,07%
40	later	29	0,07%
41	definitely	27	0,07%
42	though	24	0,06%
43	i know	21	0,05%
44	sort of/sorta	19	0,05%
45	especially	15	0,04%
46	exactly	15	0,04%
47	just like	15	0,04%
48	obviously	14	0,03%
49	mainly	13	0,03%
50	second one..	12	0,03%
51	you see	12	0,03%
52	although	11	0,03%
53	even though	11	0,03%
54	for example	11	0,03%
55	instead of	11	0,03%
56	of course	10	0,02%
57	sure	10	0,02%
58	anyway	9	0,02%
59	great	9	0,02%
60	however	9	0,02%
61	totally	9	0,02%

62	as well	8	0,02%
63	for instance	8	0,02%
64	in conclusion	8	0,02%
65	such as	7	0,02%
66	in fact	6	0,01%
67	so far	6	0,01%
68	yet	6	0,01%
69	as you (can) see	5	0,01%
70	at the same time	5	0,01%
71	extremely	5	0,01%
72	i believe	5	0,01%
73	somehow	5	0,01%
74	eventually	4	0,01%
75	in that sense	4	0,01%
76	next one	4	0,01%
77	recently	4	0,01%
78	whereas	4	0,01%
79	finally	3	0,01%
80	frequently	3	0,01%
81	in general	3	0,01%
82	initially	3	0,01%
83	quite	3	0,01%
84	right now	3	0,01%
85	simply	3	0,01%
86	afterwards	2	0,00%
87	as a result	2	0,00%
88	at the end	2	0,00%
89	briefly	2	0,00%
90	essentially	2	0,00%
91	fairly	2	0,00%
92	lemme	2	0,00%
93	specially	2	0,00%
94	specifically	2	0,00%
95	as i said before	1	0,00%
96	by the way	1	0,00%
97	considerably	1	0,00%
98	generally	1	0,00%
99	in other words	1	0,00%
100	indeed	1	0,00%
101	needless to say	1	0,00%
102	otherwise	1	0,00%
103	partially	1	0,00%
104	previously	1	0,00%
105	whatsoever	1	0,00%

TOTAL :	8.837	21,40%
----------------	--------------	---------------

In particular, the number of total words uttered by native speakers of English is 41,173 and the occurrences of discourse markers are totally 21,48 % with 8.844 hits. There are **107** different discourse markers identified with several frequency rates.

As can be observed through the table, **and** is the most recurrent discourse marker used by native speakers of English with 1519 hits by representing 3,69 % out of 21,48 %. The discourse marker **umm** follows **and** with 1232 hits by representing 2,99 %. Another mostly recurrent discourse markers are **like** with 612 occurrences (1,49 %) and **uhh** with 591 occurrences (1,44 %). Moreover, the discourse marker **so** just comes after them with 541 hits (1,31%). Thus, **and**, **umm**, **like**, **uhh** and **so** are placed at the first five of the most recurrent discourse markers used by native speakers of English.

Other significantly recurrent items can be stated as **but** with 383 hits, **just** with 367 hits, **then** with 216 hits, **okay** with 211 hits and you know with 205 hits in MICASE Corpus. Additionally, these are followed by **really**, **because**, **yeah**, **or** and **also** with similar number of occurrences.

As it was also conducted in the analysis of results within the research corpus, some items were gathered within the table; such as **mhm** and **uhuh**, **first** and **first one**, **sort of** and **sorta** and **kind of** and **kinda**, as they are the same discourse markers with slight changes in their word forms.

Furthermore, several items represent the same percentage. To give a few examples, **even**, **only** and **maybe** represent 0,17% each; or **pretty**, **stuff** and **still** represent 0,12% although each discourse marker have slightly different occurrences. Additionally, there are several items representing 0,04% (such as **exactly**, **just like**, etc.) or 0,03% (such as **mainly**, **second one**, **you see**, etc.).

As the list of discourse markers present, there are also discourse markers under the percentage of 0,03 %. The occurrences of these discourse markers are less frequent and range from 10 and 1 hits. Particularly, the discourse markers representing 0,02 % each are **anyway**, **great**, **however**, **totally**, **as well**, **for instance**, **in conclusion** and **such as**. Moreover, the discourse markers in the following; **so far**, **yet**, **as you can see**, **at the same time**, **extremely**, **i believe**, **somehow**, **eventually**, **in that sense**, **next one**, **recently**,

whereas, finally, frequently, in general, initially, quite, right now and *simply* represent only 0,01 % each in the research corpus.

Moreover, the list also presents that there are also several discourse markers that have only one hit in MICASE corpus. These are *afterwards, as a result, at the end, briefly, essentially, fairly, lemme (let me), specially, specifically, as i said before, by the way, considerably, generally, in other words, indeed, needless to say, otherwise, partially, previously* and *whatsoever*. Since these items occur just once, they are not displayed within the frequency rate.

Finally, the graph below also presents clearly top 20 most frequent discourse markers used by native speakers of English, through analysing the frequency rate of discourse markers within the total number of discourse markers, which is 8.844.

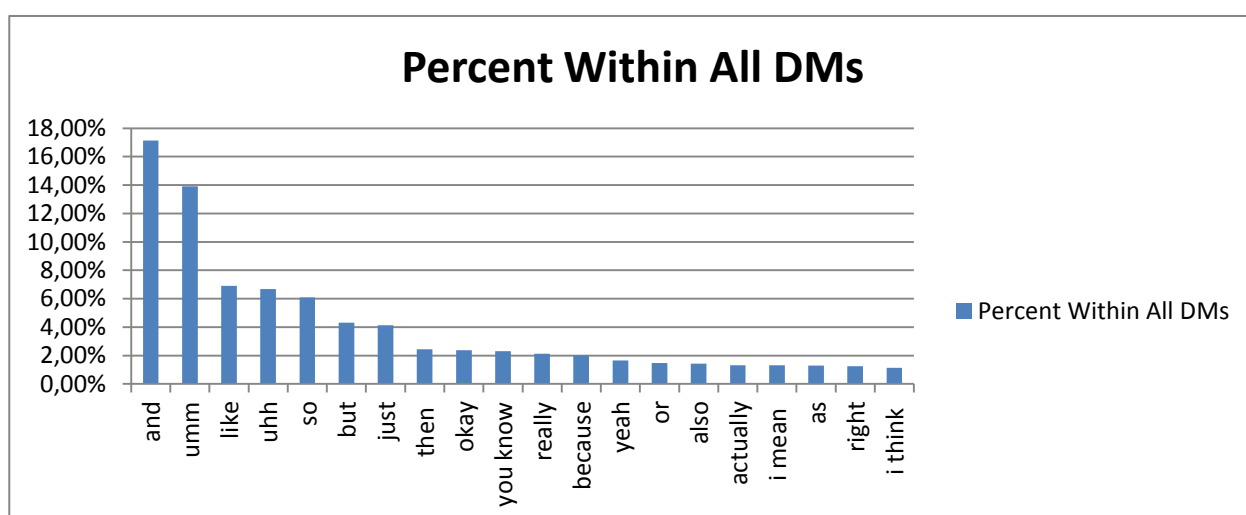


Figure 3: First 20 most frequent discourse markers in MICASE corpus

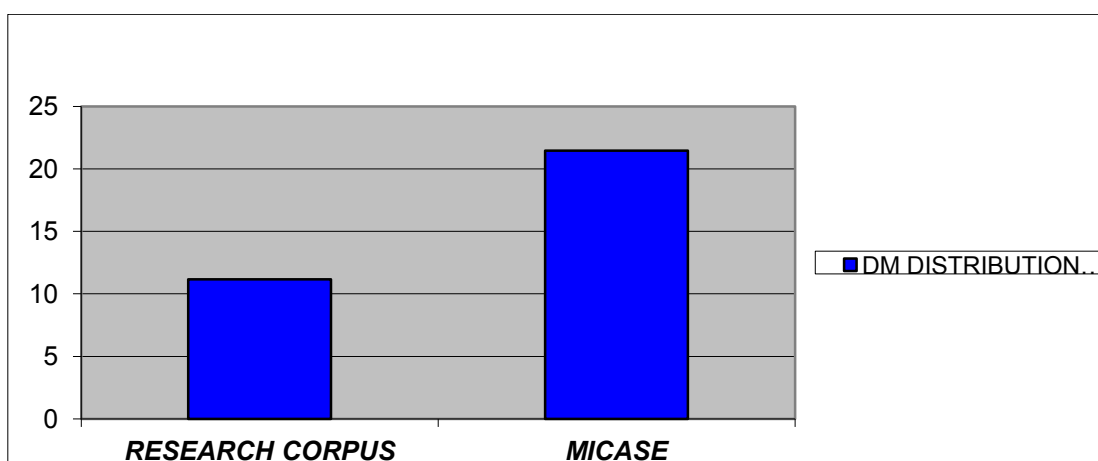
4.5. Comparative Analysis

The research provides the results of analysis of the research corpus and MICASE in the previous sections through particular tables. After all, it is necessary to compare and contrast the results of both corpus to find out the similarities or differences between them. In particular, as the research aims at searching for the comparison

between the discourse markers used by native speakers of English and Turkish non-native speakers of English, the results of both corpus have been compared.

As an overall evaluation, there have been 3.839 occurrences of discourse markers identified in the research corpus, which consists of 34.420 words. On the other hand, within 41.173 words, there have been 8.837 occurrences of discourse markers identified in MICASE. These results indicate that the frequency of discourse markers in the research corpus, 11,15 %, is lower than than the one in MICASE with 21,46 %, such as in the following graph.

Graph 3: Distribution of discourse markers in research corpus and MICASE



In addition to this overall evaluation, for the purpose of analyzing the discourse markers particularly in each corpus, one of the methods of statistics, *median*, was used. Median which can be described as the numerical value separating the higher half of a sample from the lower half is preferred to use to differentiate more frequent items from less frequent items in each particular corpus. It was also necessary to focus on the scope of the study without distracting within the myriad of discourse markers due to the complexity and variety of discourse markers in both corpus. Otherwise, dealing with each discourse marker identified in each corpus would cause disorientation in reaching the objectives of the research. Thus, medians of both corpus were determined and a comparative analysis was conducted within more frequent items of each corpus.

In the tables of results of each corpus in the previous parts, all of the items were presented. As it was also mentioned before, since some of the discourse markers has just one occurrence within the corpus, they have no contribution to the frequency analysis. Therefore, these items were excluded within the lists to be used to find median in each corpus.

For the research corpus, as stated in the following table, the median is **31**. Hence, the items between 1-31 are more frequent discourse markers within the corpus of Turkish non-native speakers of English, and the rest of the list presents less frequent items used in the research corpus.

Table 10: Frequency level of the discourse markers according to the median of the list within the research corpus

No	DM	Percent	Frequency Level
1	uhh	4,07%	More
2	and	1,85%	More
3	so	0,54%	More
4	yes	0,52%	More
5	but	0,44%	More
6	umm	0,35%	More
7	for example	0,33%	More
8	as	0,28%	More
9	because	0,26%	More
10	or	0,26%	More
11	okay	0,29%	More
12	also	0,19%	More
13	let's...	0,18%	More
14	like	0,14%	More
15	then	0,11%	More
16	hmm	0,10%	More
17	as you see	0,09%	More
18	hih, hi-huh	0,08%	More
19	just	0,08%	More
20	now	0,07%	More
21	yeah	0,07%	More
22	however	0,07%	More
23	firstly, first, first of all	0,06%	More
24	really	0,06%	More
25	actually	0,04%	More
26	of course	0,04%	More
27	such as	0,04%	More
28	you know	0,04%	More

29	clearly	0,03%	More
30	i mean	0,03%	More
31	you see, see	0,03%	More
32	alright	0,02%	Less
33	by the way	0,02%	Less
34	even	0,02%	Less
35	exactly	0,02%	Less
36	for this/that reason; for these reasons	0,02%	Less
37	i think	0,02%	Less
38	only	0,02%	Less
39	right	0,02%	Less
40	as a brief	0,02%	Less
41	generally	0,02%	Less
42	i guess	0,01%	Less
43	obviously	0,01%	Less
44	probably	0,01%	Less
45	in other words	0,01%	Less
46	well	0,01%	Less
47	as an example	0,01%	Less
48	especially	0,01%	Less
49	maybe	0,01%	Less
50	moreover	0,01%	Less
51	although	0,01%	Less
52	as a result of	0,01%	Less
53	in fact	0,01%	Less
54	in this way	0,01%	Less
55	indeed	0,01%	Less
56	just like	0,01%	Less
57	lastly	0,01%	Less
58	otherwise	0,01%	Less
59	simply	0,01%	Less
60	specifically	0,01%	Less
61	this is to say	0,01%	Less

Additionally, the median of the list of discourse markers in MICASE is **42**, as presented in the Table 11. Accordingly, the items between 1-42 are the ones that are more frequent while the others are less frequent items in the corpus of native speakers of English.

Table 11: Frequency level of the discourse markers according to the median of the list within MICASE corpus

No	DM	Percent	Frequency Level
1	and	3,69%	More
2	umm	2,99%	More
3	like	1,49%	More

4	uhh	1,44%	More
5	so	1,31%	More
6	but	0,93%	More
7	just	0,89%	More
8	then	0,52%	More
9	okay	0,51%	More
10	you know	0,50%	More
11	really	0,46%	More
12	because	0,43%	More
13	yeah	0,36%	More
14	or	0,32%	More
15	also	0,31%	More
16	kind of/kinda	0,29%	More
17	actually	0,28%	More
18	i mean	0,28%	More
19	as	0,28%	More
20	right	0,27%	More
21	i think	0,25%	More
22	basically	0,23%	More
23	now	0,20%	More
24	well	0,20%	More
25	cuz	0,18%	More
26	even	0,17%	More
27	only	0,17%	More
28	maybe	0,17%	More
29	oh	0,16%	More
30	mhm, uhuh	0,16%	More
31	alright	0,13%	More
32	pretty	0,12%	More
33	stuff	0,12%	More
34	let's...	0,10%	More
35	first, first one	0,10%	More
36	yes	0,09%	More
37	i guess	0,08%	More
38	probably	0,07%	More
39	later	0,07%	More
40	definitely	0,07%	More
41	though	0,06%	More
42	i know	0,05%	More
43	sort of/sorta	0,05%	Less
44	especially	0,04%	Less
45	exactly	0,04%	Less
46	just like	0,04%	Less
47	obviously	0,03%	Less
48	mainly	0,03%	Less
49	second one..	0,03%	Less
50	you see	0,03%	Less
51	although	0,03%	Less
52	even though	0,03%	Less
53	for example	0,03%	Less

54	instead of	0,03%	Less
55	of course	0,02%	Less
56	sure	0,02%	Less
57	anyway	0,02%	Less
58	great	0,02%	Less
59	however	0,02%	Less
60	totally	0,02%	Less
61	as well	0,02%	Less
62	for instance	0,02%	Less
63	in conclusion	0,02%	Less
64	such as	0,02%	Less
65	in fact	0,01%	Less
66	so far	0,01%	Less
67	yet	0,01%	Less
68	as you (can) see	0,01%	Less
69	at the same time	0,01%	Less
70	extremely	0,01%	Less
71	i believe	0,01%	Less
72	somehow	0,01%	Less
73	eventually	0,01%	Less
74	in that sense	0,01%	Less
75	next one	0,01%	Less
76	recently	0,01%	Less
77	whereas	0,01%	Less
78	finally	0,01%	Less
79	frequently	0,01%	Less
80	in general	0,01%	Less
81	initially	0,01%	Less
82	quite	0,01%	Less
83	simply	0,01%	Less

In accordance with the results of the median, the items that are more frequent in both corpus and their frequencies are presented with the following table.

Table 12: Comparative results of the frequencies in the research corpus and MICASE corpus

Discourse Marker	MICASE	Research Corpus	Representation of DMs in research corpus as compared with MICASE
and	3,69%	1,85%	Less frequent
umm	2,99%	0,35%	Less frequent
like	1,49%	0,14%	Less frequent
uhh	1,44%	4,07%	More frequent
so	1,31%	0,54%	Less frequent
but	0,93%	0,44%	Less frequent

just	0,89%	0,08%	Less frequent
then	0,52%	0,11%	Less frequent
okay	0,51%	0,29%	Less frequent
you know	0,50%	0,04%	Less frequent
really	0,46%	0,06%	Less frequent
because	0,43%	0,26%	Less frequent
yeah	0,36%	0,07%	Less frequent
or	0,32%	0,26%	Less frequent
also	0,31%	0,19%	Less frequent
actually	0,28%	0,04%	Less frequent
i mean	0,28%	0,03%	Less frequent
as	0,28%	0,28%	Comparable
now	0,20%	0,07%	Less frequent
let's...	0,10%	0,18%	More frequent
first, first one, firstly, first of all	0,10%	0,06%	Less frequent
yes	0,09%	0,52%	More frequent

Thereupon, the more frequent items in each corpus are matched, and the result indicates that there are 22 items identified as the same in each corpus. However, their frequencies in each particular corpus indicate discrepancies. The table above also presents the representation of discourse markers in research corpus when compared with MICASE. In other words, it indicates whether the discourse markers of non-native speakers of English are used more frequently or less frequently by comparing with the same discourse markers used by native speakers of English.

As a consequence, 18 discourse markers (for instance *and*, *umm*, *you know*, *okay*, *just*, *actually* etc.) out of 22 are used more frequently by native speakers of English while only three discourse markers (*uhh*, *let's...* and *yes*) are used more frequently by non-native speakers. Besides, *as* is the only one which has the same frequency in both corpus.

Moreover, the other discourse markers in the lists should be taken into account as they also reflect significant discrepancies. For example, the discourse markers in MICASE such as *kind of/kinda*, *right*, *i think*, *basically*, *well*, *cuz*, *even*, *only*, *maybe*, *oh*, *alright*, *pretty*, *i guess*, *probably*, *later*, *definitely*, *though* and *i know* do not exist in the more frequency table of the research corpus. On the other hand, there are also some

items of the research corpus that do not exist in more frequent table of MICASE, which are *for example, as you see, however, of course, such as, clearly, you see/see*.

4.6. Analysis of Functions of Discourse Markers in the Research Corpus and MICASE Corpus

The research aims at not only providing quantitative analysis of the discourse markers used by Turkish non-native speakers of English but also presenting how the identified discourse markers are used. Thus, qualitative side of the analysis consists of dealing with functions of discourse markers through some examples. As the nature of discourse analysis suggests, functions of items are highly fundamental in analysis. It was aimed through qualitative analysis to enrich the research with concrete examples of discourse markers from both the research corpus and MICASE.

This part of the research is based on Fung and Carter's (2007) multi-categorical framework of discourse markers. By following Maschler (1994, 1998), Andersen (2001), Lenk (1998), Jucker and Ziv (1998), their framework includes discourse markers that can be categorized under four functional headings although any instance may perform more than one of these functions. However, it should be noted that due to the complexity and variety of the functions of discourse markers, this table is taken as the basis to take the framework as limitation for the purpose of giving examples of discourse markers.

Table 13: A core functional paradigm of discourse markers in pedagogic discourse

Interpersonal	Referential	Structural	Cognitive
Marking shared knowledge: <i>See, you see, you know, listen</i>	Cause: <i>Because, cos</i> Contrast: <i>But, and, yet, however, nevertheless</i>	Opening and closing of topics: <i>Now, OK/okay, right/alright, well, let's start, let's discuss, let me conclude discussion</i>	Denoting thinking process: <i>Well, I think, I see, and Reformulation/Self-correction: I mean, that is, in other words, what I mean</i>
Indicating attitudes: <i>well, really, I think, obviously,</i>	Coordination: <i>And</i> Disjunction: <i>Or</i>		

<i>absolutely,</i>			<i>is, to put it in</i>
<i>basically, actually,</i>	Consequence: <i>So</i>	Sequence: <i>First,</i>	<i>another way</i>
<i>exactly, sort of,</i>		<i>firstly, second,</i>	
<i>kind of, like, to be</i>	Digression: <i>Anyway</i>	<i>secondly, next,</i>	Elaboration: <i>Like, I</i>
<i>frank, to be honest,</i>		<i>then, finally</i>	<i>mean</i>
<i>just, oh</i>	Comparison:		
	<i>Likewise, similarly</i>	Topic shifts: <i>So,</i>	Hesitation: <i>Well, sort</i>
Showing responses:		<i>now, well, and of</i>	
<i>OK/okay, oh,</i>		<i>what about, how</i>	
<i>right/alright, yeah,</i>		<i>about</i>	Assessment of the
<i>yes, I see, great, oh</i>			listener's knowledge
<i>great, sure, yeah</i>		Summarizing	about the utterances:
		opinions: <i>So</i>	<i>You know</i>
		Continuation of	
		topics: <i>Yeah, and,</i>	
		<i>cos, so</i>	

(Fung and Carter, 2007:418)

The examples below are given from both the research corpus and MICASE to display descriptive details in parallel with the categories and subcategories of the table above.

4.6.1 Interpersonal Category

Discourse markers have a pivotal role in increasing interpersonal attitudes between speakers in spoken discourse. They are used to mark shared knowledge between speakers. For example, “by using *you know* the speaker wants the hearer and accept the content of the speaker’s utterance, without arguing; so it functions as a face-saving device by avoiding the possible opposition to be made by the hearer” (Östman, 1981, p.17).

..., the people who did, come in with different types of responses, were the native speakers, and none of the non-native speakers

thought to even accept it or avoid it. they they basically had, **you know** three types of responses that they could draw from, whereas the native speakers have a much wider... (MICASE)

The phrase **you know** in line 3 functions as a discourse marker in medial position within the sentence. In the extract, it is used to include the listener to the issues following, specifically, types of responses so as to state that they can share some kind of knowledge together. Fox Tree and Schrock (2002) also points out that “**you know** may be more common in talking to friends than strangers because friends share more mutual knowledge; that is, speakers may be more likely to invite addressee inferences when they believe inferences drawn will approximate their thinking” (p.737). Since the speakers in the corpus are classmates, it is very usual to use **you know** in presentations.

Another sub-category of interpersonal functions is indicating attitudes. In particular, within this category, many adverbs such as really, basically, actually, etc. are used commonly. In the extracts below, **actually** and **basically** are used generally in initial position of the utterances following another utterance to show the speaker’s own attitude about the statement.

... John, John was mean and his wife was even meaner instead of in one means indirectly. Austens knows_, **actually** her novels are umm often read for the entertaining way in which they are written rather than the plot (RESEARCH CORPUS)

... so **basically** most of them did not grow up in a bilingual community, except for that of their parents, and i've done qualitative interviews i haven't, finished evaluating all of that yet but **basically**, for the qualitative, work it seems like most of the students, did indeed grow up in um an integrated community, and they often say well i spoke Spanish or Hindi with my parents friends... (MICASE)

Fung and Carter (2007) states that **kind of** and **sort of** also function “to reduce the face-threatening act by having the interactive effect of softening the tone through an element of vagueness” (p.417), as in the following extract.

... and you can see that, um, this is, these are both for the Spanish sample. so for the Spanish sample i mean these are pretty low numbers so it's **kind of** weird to see it in this form but this just basically means that, we are (xx) (black) (line) (notes) fifty percent, whatever answered higher than that, so you can see that everybody, who's yes involve (MICASE)

Moreover, “*I’m not saying... I’m just saying*’ is a structure frequently used to hedge a standpoint against actual or anticipated criticism while simultaneously asserting that the standpoint has been essentially continuous and remains unchanged” (Craig and Sanusi, 2000, p.434). Similarly, **just** in the following example emphasizes this standpoint continuity of the argument.

S5: so she had it planned for next week already... [S1: this poor woman] we've edited out all the names in the email... okay, so we **just** thought that um, those two things we didn't take into consideration that the content is much_ would be overlapping of pragmatic and linguistic knowledge but we thought, that those were, possible,.. (MICASE)

The collocation **oh really** also functions as discourse marker although **oh** and **really** function separately as discourse markers indicating attitudes. In particular, Castro (2009) states that “**oh really**, with upward intonation, is both expressing a response (of surprise) and requesting confirmation from the students” (p. 72), as in the following extract from MICASE. Additionally, there is no occurrence of this collocation in the research corpus.

S12: hello everybody... um how many of, you have heard of Ruby Dee...? **oh really** that's it? okay um, how many of, you have seen Do the Right Thing...? Jungle Fever...? and more recently Baby Geniuses...? <LAUGH> okay well, um i recently wrote an article about Ruby Dee,

Another interpersonal type of discourse marker is the discourse marker group that is used to show responses such as *okay, right/alright, yeah, yes, I see, great* and *sure*.

S1:..., when do you use indirectness uhh in your daily lives? when do you use? or in which situations? to whom do you use indirectness or directness? can you give me an example?

S2: the situation?

S1: *yes*.

S3: for example, uhh when, when i try to tell one of my friend is sad uhh. is the (xx) when she uhh wears something, for example, i don't want to say, it's not good. (RESEARCH CORPUS)

Yes is used more frequently by the students in the research corpus as a discourse marker to show responses. However, native speakers use a variety of discourse markers to indicate their responses, such as *yeah, right* and *sure*, as in the following extracts. Especially, the token *yeah* is among the most common discourse markers of native speakers that can function in various ways such as “back-channel cue, acknowledgement token, continuer and speaker incipency” (Wong, 2000:41).

S5: well we talked about that a little bit didn't we and i think we decided that it didn't constrain very much?

S3: *yeah* because i i was saying with my Spanish i don't know much Spanish, but i could just as easily say, i, have plans from eleven to one so i can't be there, as i could say i'm really busy this week,... (MICASE)

...but in terms of the academic stuff, it would seem to me that the class actually encour- encourages the students to look at Spanish in a new domain. [S1: *right*] encourages very academically based [S1: *right*] so it would be interesting to actually, perhaps, in another time another [S1: mhm] place, [S3: a new world] <SS: LAUGH> to actually see you know...(MICASE)

and this person says i say i probably wouldn't respond. so, that's, the point he wouldn't respond that's what the point is...

S5: can i just clarify something?

SU-f: **sure** go ahead

S5: **sure**. um i just wanted to clarify that, s- sorry, when we began, we didn't anticipate all of these categories, we didn't anticipate acceptance, uhh the email was worded it didn't say that.. (MICASE)

4.6.2. Referential Category

The second category for functions of discourse markers is the referential level. Referential level includes the discourse markers that are used to make references between utterances, particularly conjunctions. Conjunctions like *because*, *although*, *however* are very recurrent in written discourse to indicate the relationships between an existing statement with the preceding one. Thus, they are similarly used in spoken discourse to display the relationships between utterances.

The discourse markers within referential level are classified into several groups; causal (*cos/because*), consequential (*so*), contrastive (*but*), disjunctive (*or*), coordinative (*and*), digressive (*anyway*) and comparative (*likewise, similarly*).

Within the sub-category of causal discourse markers, *because* and *cos/cuz* takes place. As in the following extract from the research corpus *because* is used to state the reason between why someone feels close to another person.

and things like these which together determine the overall degree of respectfullness within a given speech situation. in other words, if you close uhh if you feel close to someone **because** that person is related to you, or you know him or her well uhh or he or she is similar to you in terms of your age, social class, occupation etc. (RESEARCH CORPUS)

Cuz/cos which can be described as phonological reduction of *because* (from *because* to *cause* and then to *cuz/cos*) is frequently used by native speakers by referring the same discorsal effect as *because*. Native speakers prefer using *cuz* as it is easy to handle it within the flow of their conversation and it serves too much to their fluency, as given in the extract below. Moreover, within the analysis of the research corpus, any occurrence of *cuz* is not determined.

um that i could get thrown into that classroom, um... and... i don't know i was kinda, it kinda intimidates me *cuz* i don't know what i would do, you know if i was in her place, um *cuz* i don't know karate. <SS: LAUGH> i asked myself you know why i'm so scared of this classroom, um. (MICASE)

On referential level, Schiffirin (1987) explains the significance of *but* as “That *but* is an adversative conjunction suggests that what follows *but* is an idea which contrasts with what has preceded” (p.52). *But* is frequently used by native and non-native speakers of English to state the contrasting points, as in the following extract:

...the speaker says something else and the umm listener, interlocutors understand something different *but* not uhh he feels himself that_ uhh the speaker is saying something *but* he says the thing not directly, indirectly. (RESEARCH CORPUS)

So which is another multi-functional discourse marker is used to state the consequential relationship between two subsequent utterances. Moreover, Fraser (1990) refuses to see “*so* solely as a marker of result or consequence and points that *so* as a discourse marker permits a range of interpretations” (p.393). For the consequential relationship of *so*, the following extract illustrates that the speaker states the result of being close to the people by connecting the utterances with *so*.

...they are so often confused is that they refer very frequently co-occur, we tempt to be socially distant from those in power over us. because uhh we are not close to those people, *so* we use indirectness. but this is not always the case. uhh somebody clearly shows how in the language classroom students are often

so close to their teachers even though there is a mark...
(RESEARCH CORPUS)

Another significant discourse marker on referential level is *anyway* which functions as leaving a point behind as in the following extract. Coll (2009) specifies this dismissive function of *anyway* as “this use always connects to a previous piece of discourse, as the previous piece of discourse is what is being dismissed, or considered an unimportant matter” (p.181).

...like when we talk about how, like, television shows that reflect black life don't, accurately reflect black life that television is not an accurate representation of life *anyway* so i think that that was a very important point that you, brought up a moment ago.

SU-f: but i think (MICASE)

4.6.3. Structural Category

Discourse markers have also functions to make successive units of talk organized. In particular, opening or closing the topics, transition between the utterances, continuation of topics and summarizing them can be conducted by using certain discourse markers.

First group under structural category of discourse markers is the ones serving to opening new topics or closing the existing ones, which are *now, okay, firght, well, let's start, let's discuss*. Especially, within the research corpus, non-native speakers start their presentations or open new topics by using *okay* or starting the phrases with *let's*, which is more frequent in research corpus. The extract below illustrates how the speaker wants to take action by using specifically *let's* in the initial position of the utterances.

Let's look at the headlines, uhh intentional indirectness, uhh indirectness is costly and risky uhh assumption of rationality, the principle of expressibility and and an illustration. *okay, let's* begin with intentional indirectness. (RESEARCH CORPUS)

As for native speakers, *well* is among the most recurrent discourse markers used to open topics. In fact, *well* takes place in the category of discourse markers that have several and various functions in spoken discourse. Svartvik (1980) summarises his findings about *well* as a qualifier to indicate or mark agreement, positive reaction or attitude, reinforcement, the non-straight and incomplete answer to the wh-question, a non-direct or qualified answer and as a frame, normally occurring non-initially, *well* shifts the topic focus to one of the topics which have already been under discussion, introduces explanations, clarifications, etc., indicates “the beginning of direct speech”, functions as “editing marker for self-correction”. Thus, *well* is commonly used by native speakers of English for several purposes. The following extract exemplifies its function of opening the presentation and its use in initial position. Besides, *well* is determined as significantly less frequent in the research corpus.

S4: um *well* thank you for being present for our presentation and um, um our project is about, is a combination of cross-sectional, studies and pragmatics, and well through this class we all know that there are, (MICASE)

In terms of narrating the topics in a sequence, *first*, *firstly*, *next*, *and then* or *finally* are commonly used as discourse markers, as in the following extract.

first of all, we have three questions about the three conversations. please read it *and then* uhh read the conversations. which one involves two teenage girls, which one involves an adult and a toddler and which one involves an elderly person and a young, younger adult. (RESEARCH CORPUS)

One of the essential roles of discourse markers is shifting the existing topic to another one. *So*, *now*, and *well* are used for that purpose. The extract below illustrates how *so* and *now* used sequentially to “signal the transition of a topic, marking the end of a topic and the beginning of another one” (Carter, 2002:421).

but i'm mainly going to be talking about is how it causes elevated levels of cortisol in your body and um, how that creates many problems if you have too much cortisol if you

stress too much. **so now** i'm gonna talk about cortisol that's just uh, the structure of it. um cortisol is also known as hydrocortisone, it's uh it's the principal glucocorticoid and glucocorticoids um, (MICASE)

To specify that the conversation has come to an end and summarize the opinions, again **so** can be used. In that occurrence, generally **so** takes initial position of the utterance, as illustrated below.

for example, woman women say atashi vatashi. [SS: LAUGHTER] vatakushi the most formal one. but men or e and bugu. bogu <LAUGHS> the most casual one. **so** let's come to the conclusion. gender exclusives speech form reflects social status or power, power differences and gender exclusive social roles. (RESEARCH CORPUS)

On structural level, discourse markers such as **yeah**, **and**, **cos**, and **so** are used as continuers to indicate the intention of the speaker to hold the floor. The following extracts include various discourse markers used by the speaker to display that s/he is still talking and will talk for a while.

and then she would email and say some plans came up <SS: LAUGH> **yeah and so** that would be avoidance, that was one of the people who avoided responding or avoided rejecting someone, **so, yeah** that that was interesting. (MICASE)

4.6.4. Cognitive Category

One of the fundamental contributions of discourse markers to spoken discourse is that they enable cognitive processes of the speakers to be observed. In other words, they provide clues about how the speaker is thinking at that specific moment, how he or she shall organize his or her thinking process or possible what shall come after the existing utterances. Hence, this characteristic of discourse markers are named as cognitive category.

Within cognitive category, the function of discourse markers to denote the thinking process are fulfilled by *well*, *i think*, *i see* and *and*. *Umm* and *uhh* is also used for the same purpose, and they are significantly recurrent in the research corpus and MICASE. The following extracts from both corpus (although *well* has limited occurrences in the research corpus) illustrates the use of *well* as the speaker thinks what he or she will say after a period of pause and *and um* for the same purpose.

we uh tried to make a chart out of, we tried to categorize the responses, and um we made, six, categories, and... and... we tried to um, define the terms... by ourselves, and <PAUSE:19> um... **well** first you see, there are six (not yet,) six categories, and the first one is acceptance and, we s- defined it as accepting the offer even though you don't want to, (RESEARCH CORPUS)

one that concerns the professor, and to two of them we, sent, the, one that concerns the classmate and the other one, th- the rest of them we sent, both of the scenarios. **and um**, that was because um, Professor Johnson said that, if we sent two of them to all of them they would be aware about the status differences and they would, kind of, response in the way they should res (MICASE)

Moreover, some of the discourse markers are specifically used to reformulate the utterances. In written discourse one can have the chance to rewrite the sentences after thinking for a while; on the other hand, as Fung and Carter (2007) supports, “speakers in real speech are under time constraints to structure and formulate their ideas” (p. 424). Thus, discourse markers are used to reformulate, rephrase, self-correct or repair their utterances. For that purpose, *i mean*, *in other words* or *that is* are used, as illustrated below.

and things like these which together determine the overall degree of respectfulness within a given speech situation. **in other words**, if you close uhh if you feel close to someone because that person is related to you, or you know him or her

well uhh or he or she is similar to you in terms of your age, social class, occupation etc. (Research)

Additionally, particular discourse markers such as *i mean* and *like* are used to elaborate the existing propositional meaning to clarify the intention of the speaking or to support the utterance. For instance, Schifffrin (1987) points that I mean is used to modify the speaker's own idead and intentions, as in the following example.

S3: even allowing for that though i think there's more more nouns are transferred [S7: oh] than in the language as a whole [S7: yeah it's not] there're there are syntactic re- well. *i mean* basically the argument is they don't drag a lot of syntax around with them <SS: LAUGH> you know if you have a verb for example some verbs are transitive some intransitive, [S4: Speaker information re (MICASE)

Moreover, *like* as illustrated below is most frequently used by native speakers of English to exemplify the previous units of talk and elaborate the utterance.

... like toward the reference area and then i came across a debate i wasn't even aware that they had this but they have like, um set of curriculum, um, *like* the standard curriculums for *like* science and math for the Ann Arbor, Public Schools so i looked through them, and then, i looked under like, um the elementary grades of, and it was like required like, they listed the science topics (MICASE)

Another significant function of discourse marker is that they are used as hesitation markers, like *well* and *sort of*, which is less frequently used by non-native speakers in the research corpus. They used *uhh* more frequently for stating their hesitations during the speech. The following extract presents how *well* is used by the speaker when he or she hesitates for a moment, which is a very recurrent moment for native speakers.

and then in parentheses he goes until the fifteenth in the olden times which coulda been all of twenty years ago for all we know <SU-f: LAUGH> and the period referred to as, uh *well* um, i can't say this it's in Japanese but in parentheses he goes cash is paid to children for New Year presents by their parents and grand- grandparents, children totally get a hundred to five hundred (MICASE)

In summary, discourse markers can serve a wide range of functions in spoken discourse from summarizing topics to denoting thinking process, from elaborating the ideas to showing responses. Particularly, the extracts above illustrate these various functions of discourse marker. Moreover, they also show that how a certain discourse marker can be used for different purposes like *so* and *well* and how they can be more or less frequent in each corpus.

4.7. Discussion

The results of the data analysis are presented in the previous sections through tables and examples from both the research corpus and MICASE. This section discusses the overall findings of the research. Through the quantitative and qualitative analysis, the research reaches its objectives, which are, identifying the discourse markers and their frequencies of Turkish non-native speakers of English and native speakers and then conducting a comparative analysis through the results. Thus, the findings are analyzed to reach certain comments about the discourse markers used by both non-native speakers and native speakers of English.

One of the points to be mentioned is about the corpus descriptions as mentioned in the Methodology section through tables displaying the titles, subject areas, primary discourse modes, duration and word count. According to this description and analysis of the transcripts of Turkish students (see Appendix C), it is noted that students' presentations of Turkish non-native speakers of English are mostly carrying the features of monologic discourse mode whereas the presentations of the students in University of Michigan offer much more interaction in their presentations. The reason underlying the difference may be that native speakers feel much more relaxed in their native language

as they are born with and live through with English and they live numerous interactions in several domains with English. However, the domain for Turkish students is highly limited to the courses that they are attending in their department. Thus, it is obvious that they may feel much less relaxed while presenting a topic in a foreign language. In particular, the interactivity rating of the presentations play a significant role in identifying and discussing the discourse markers of the students in spoken discourse. Discourse modes of the presentations effect the types of discourse marker to be used by the speakers. Thus, it can be concluded that the monologic characteristic of the presentations of Turkish students have influenced their use of discourse markers in a way that they use more structural and referential discourse markers rather than interpersonal ones in their presentations.

As an overall evaluation of the occurrences of the discourse markers between two groups of speakers, the findings of the study show that native speakers use discourse markers more frequently than non-native speakers in terms of occurrences within the total word count of the transcripts and also native speakers use much more different discourse markers with several functions, that is, their spoken discourse has a variety of discourse markers. This overall finding reflects the outcomes of the previous studies by Weinert (1998), Trillo (2002) and Hellerman and Vergun (2007). However, it cannot be stated that non-native speakers within the study have used discourse markers in limited number. In fact, it can be argued that there is a *tendency* to use discourse markers in their presentations. This result also supports the previous studies on the discourse markers by Hays (1992), Lee (1999) and Hellermann & Vergun (2007). They claim that students with a higher proficiency in the learned language who are more acculturated to the L2 environment are more likely to use discourse markers. Although the participants of the current research corpus are not acculturated to the the foreign language environment, they are upper level students of English and their use of discourse markers is significant, which supports Hellerman and Vergun's (2007) statement that more proficient students in the upper level classes use more of the focal discourse markers. However, the occurrences of discourse markers of Turkish non-native speakers are acceptable, it is resulted that it is not satisfying when compared to the results of the analysis of native speakers and when analyzed according to several functions of discourse markers. Thus, this tendency of using discourse markers in spoken discourse

should be supported to make Turkish non-native speakers fluent in their spoken discourse.

As an overall evaluation of quantitative analysis of the current research, it can be noted that the most frequent discourse markers used by native speakers of English such as *and, like, so, but, just, then, okay, you know, really, yeah, i mean* etc. exist in the table of the most recurrent items used by Turkish non-native speakers. However, when compared with native speakers, these tokens are less frequently used by non-native speakers. Moreover, particular discourse markers such as *kind of/kinda, right, i think, basically, well* and *cuz* do not take place in the more frequent items of non-native speakers although they are considerably used by native speakers.

One of the fundamental points to be discussed is that the most frequent item of Turkish non-native speakers is *uhh*. In particular, it represents 4,07 percent of the whole corpus. This frequency (see Graph 1) is considerably significant when compared to the frequencies of other discourse markers in the research corpus. When this item is excluded from the analysis, total frequency of discourse markers in the research corpus becomes 7,08 percent, which is highly lower. Moreover, Turkish non-native speakers of English use *uhh* for several purposes such as hesitation, denoting thinking process, searching for the right word and fillers. Therefore, it can be concluded that *uhh* is a ‘savior’ for Turkish non-native speakers and instead of using any kind of discourse markers they utter *uhh* most of the time. Thus, this result clearly supports the argument of the research and the main reason of overusing and fossilization of *uhh* is that Turkish non-native speakers are not aware of and competent in using particular discourse markers for particular contexts and several functions. Therefore, there is definitely a need to learn more types of discourse markers and how to use them.

When comparing the results of the two corpus, it can also be observed that Turkish non-native speakers use *uhh* more frequently while native speakers use *umm* more frequently. Clark and Fox Tree (2002) point out with a corpus study on native speakers of English that “speakers use *uh* and *um* to announce that they are initiating what they expect to be a minor (*uh*), or major (*um*), delay in speaking to implicate that they are searching a word, deciding what to say next and want to keep or cede the floor” (p.73). Similarly, Turkish non-native speakers use *uhh* for specific purposes.

Particularly, this comparison may also reveal a certain sociolinguistic difference between native speakers and Turkish non-native speakers. *Uhh* is also commonly used by Turkish speakers when they are also talking in their mother tongue. Hence, they transfer this particular feature of their spoken discourse in Turkish to English. On the other hand, native speakers of English use *umm* for different purposes in their conversations. Thus, *umm* can be specified as culturally specific for native speakers of English.

Another significant observation is the use of *yes* and *yeah* by the speakers. *Yes*, formal form of *yeah*, is one of the few discourse markers that is more frequent in non-native corpus. On the other hand, *yeah* is a more frequent item used by native speakers of English. This finding also supports the result of Fung and Carter's (2007) research in which they are comparing the frequencies of native speakers and Hong Kong language learners. Fung and Carter (2007) also conclude that "there is an over reliance on *yes* rather than *yeah* among the Hong-Kong subjects and they did not use the range of possibilities available with *yeah* that native speakers do as a way to exhibit understanding or acknowledgement (interpersonal category), or as a continuer of the progress of the primary speaker's turn (structural category)" (p.431). Moreover, native speakers use different types of discourse markers to show responses like *yeah* such as *sure*, *right*. On the other hand, again non-native speakers do not have this kind of variety.

Similarly, non-native speakers use *because* for referential purposes while native speakers again use both *because* and *cuz*. However, non-native speakers do not use *cuz* at any instances, but *cuz* functions in a more overtly discourse-marking role in spoken discourse rather than *because* as it serves continuation of topics as well as causal relationships between utterances.

Moreover, the case of *well* should also be noted. Although *well* is used by native speakers for several purposes, non-native speakers do not benefit from several functions of *well* that can be used for various instances in their spoken discourse. Since *well* is the most frequently analysed discourse marker by different authors (e.g. Cuenca, 2008; Blakemore, 2002; Aijmer and Simon-Vandenberg, 2003) due to its significance in spoken discourse by being fully pragmatic in terms of structural, cognitive and

interpersonal functions, it is notable that non-native speakers in the research corpus are not particularly capable of using *well* in their conversations.

To put it in a nutshell, the findings display that discourse markers are not totally excluded in the research data but they are used less frequently. In other words, Turkish non-native speakers tend to use less frequently the kind of discourse markers that native speakers use mainly in interpersonal and cognitive categories. On the other hand, the discourse markers identified in the research corpus are primarily in the textual category, with a significantly heavy use of referential and structural tokens. Therefore, the variety and the range of discourse markers used by Turkish non-native speakers are limited in and confined to particular items, and thus there is an overreliance on certain discourse markers which lead to pragmatic fossilization. Moreover, as Qun (2009) highlights with illustration of “and so on”, in cross-cultural communication that this kind of overused and fossilized items avoid the use of other types of discourse markers or mobilized so as to add variety to language. Hence, Turkish non-native speakers of English should be aware of the discourse-pragmatic competence that several discourse markers may serve.

CHAPTER 5

CONCLUSION

5.1. Introduction

This chapter presents an overall review of the current research including brief summaries of the background, objectives, methodology, data collection procedure and results of analysis. Then, these major parts of the research are followed by the pedagogical implications that should be mentioned so as to provide specific suggestions about discourse markers in teaching English as a foreign language. Finally, implications for further research are highlighted to propose new dimensions within the analysis of discourse markers in the field.

5.2. Summary

In recent years, it has been accepted that teaching and learning foreign languages require much more than focusing on only surface grammatical and lexical systems of the language. Having communicative and pragmatic competence in target language has gained impetus according to the requirements of the new era. Nowadays, language learners need the ability to use lexical items of the target language functionally according to appropriate context in appropriate domains. In parallel with these advancements, discourse analysis has attained a pivotal role in providing language teaching more concrete examples about the units of language. With the help of discourse analysis, the fundamental elements within written discourse and spoken discourse have been investigated. Within spoken discourse, there are some fundamental units of talk that serve several functions. Discourse markers are considered among the essential elements in spoken English discourse as they function for multiple purposes and activate discourse-pragmatic competence of the speakers.

This study aimed at identifying the discourse markers used by Turkish nonnative speakers of English. The questions guiding the research are: which discourse markers

are used by Turkish nonnative speakers and in which frequency levels, whether the results differ from the ones of the discourse markers used by native speakers of English and what kind of functions of discourse markers are used by native and nonnative speakers.

In order to reach the objectives of the study, a corpus-based analysis was implemented. This type of analysis deals with the corpus in question quantitatively and qualitatively. To this end, at first, the characteristics of the corpora such as setting, participants and data collection instruments were decided. For the corpus which includes the spoken transcripts of Turkish nonnative speakers of English, thirty presentations of senior-undergraduate students studying in Department of English Language Teaching in Gazi University were audio-recorded and according to audio quality and the representation level of the discourse, twenty presentations were selected to be transcribed. Each selected presentation was transcribed in accordance with MICASE Transcription Conventions. Meanwhile, to conduct a comparative analysis, transcripts of the presentations of senior-undergraduate students studying in University of Michigan were obtained from MICASE online corpus.

After the transcription process, for the quantitative side of the analysis, the frequency analysis was conducted according to the occurrences of discourse markers used in the transcripts. With the help of AntConc programme, concordance lines for each discourse markers were arranged. Each instance of a particular discourse marker was analyzed whether it fulfills the functions of discourse markers or not. Thus, the frequency analysis of the items were determined. In order to compare the results of both corpus, the median tables were composed. 31 items from research corpus and 42 items from MICASE were determined as more frequent discourse markers in their own corpus. Then, the same discourse markers were identified within more frequent tables of both corpus. Finally, 22 discourse markers were matched in both corpus so that their frequencies could be compared.

Moreover, for the qualitative side of the analysis, a core functional paradigm for the functions of discourse markers (Fung and Carter, 2007) was taken as the basis. The functions of the discourse markers were analysed according to the four categories included in this paradigm which are interpersonal, referential, structural and cognitive.

Concrete examples were given from research corpus and MICASE in accordance with these four categories in the section of qualitative analysis.

The findings of the quantitative and qualitative analysis of the study are summarized as follows:

- The use of discourse markers in spoken discourse is not totally excluded by Turkish nonnative speakers; in fact, they have remarkable tendency to use discourse markers. However, when compared with the results of nonnative speakers, Turkish nonnative speakers use discourse markers in their speech much less frequently than native speakers of English.
- Turkish nonnative speakers do not use a variety of discourse markers in their spoken discourse although native speakers prefer using a range of discourse markers in their speech.
- Turkish nonnative speakers of English overuse certain discourse markers (*uhh* initially, *yes*, *and*, *but*, *so*,) so that other types of discourse markers could not be activated to add variety to their speech and this also leads to fossilization of the discourse markers, which finally cause lack of discourse-pragmatic competence.
- The most frequent discourse markers used by native speakers of English such as *and*, *like*, *so*, *but*, *just*, *then*, *okay*, *you know*, *really*, *yeah*, *i mean* etc. exist in Turkish nonnative speakers' speech less frequently.
- Specific discourse markers such as *kind of/kinda*, *right*, *i think*, *basically*, *well* and *cuz* have significantly less frequent use by Turkish nonnative speakers in spite of the fact that native speakers often use these items considerably in their speech.
- The discourse markers determined in spoken discourse of Turkish nonnative speakers are primarily in the textual category including notable use of

referential and structural tokens while native speakers use discourse markers serving interpersonal and cognitive purposes.

Along with the findings above, this study presents the detailed comparative analysis of the discourse markers used by Turkish non-native speakers of English and native speakers of English.

5.3. Pedagogical Implications

It is widely acknowledged that second or foreign language learners should have communicative and pragmatic competence within the target language to be an effective and competent users of the language. For this specific purpose, language learners should be exposed to several functions of the elements within spoken discourse. Thus, as Müller (2005) also supports, discourse markers contribute to the pragmatic meaning of utterances and play an important role in the pragmatic competence of the speaker. In addition, it is also highlighted by Schifffrin (2001) emphasizing that discourse markers reflect not only about the linguistic properties (semantic and pragmatic meanings and functions) and the organization of social interactions, but also about the cognitive, expressive, social and textual competence of those who use them.

Native speakers of English have discourse-pragmatic competence through the use of discourse markers for several purposes in their interactions, which makes learning the use of discourse markers for non-native speakers of English pedagogically significant. Thus, the current study has also put forward this pedagogical implication about the necessity to integrate the use of discourse markers in language teaching environment, in agreement with Fung and Carter (2007) who highlight, “the restricted range of discourse markers used and the frequency of particular markers reflect the unnatural linguistic input ESL learners are exposed to and the traditional grammar-centred pedagogic focus which has been geared towards the literal or propositional (semantic) meanings of words rather than their pragmatic use in spoken language”.

As a result, there is a need for “structural and functional description of discourse markers” (p.3) as Aijmer (2002) suggests. Moreover, as cited in Müller (2005), several authors agree that lack of or wrongly used discourse markers in spoken discourse not only cause misunderstandings but also cause negative judgements on the nonnative

speaker (Svartvik 1980:171; Erman 1987:1; Nikula 1993:127) and a “giveaway” of the speaker’s “foreignness” (Hasselgren 2002:103). Wierzbicka (1991) also highlights the significance of discourse markers which underlies one’s ability to use language in culturally, socially, situationally appropriate ways which makes language learners maintaining discourse cohesiveness, communicative effectiveness and also interpersonal and cross-cultural interaction.

As this study reveals, the use of discourse markers is neglected in pedagogic settings for nonnative speakers. However, there are certain ways that can be suggested to teach discourse markers in language classrooms. Since the study displays that Turkish nonnative speakers of English are not capable of using more discourse markers effectively in their spoken discourse, using several functions of discourse markers like interpersonal and cognitive functions and using them in variety, their awareness should be raised towards the variety and functions of discourse markers (see Appendix E for sample activities).

Discourse markers can be taught by both explicit and implicit teaching. Rose and Kasper (2001) suggest that explicit instruction of the target language pragmatic rules in language classrooms raises pragmatic awareness effectively while McCarthy (1998) highlights that production should be delayed until suitable natural opportunities arise in an implicit way (see Appendix E1). In particular, integration of several activities is suggested like Fung and Carter (2007) by stating that “the language awareness-based III (Illustration-Interaction-Induction) approach proposed by McCarthy and Carter (1995), mediated through activities like language observation, problem-solving, and cross-language comparisons, can be illuminating in bringing out the meaning and usage of various discourse markers” (p.434). Moreover, as Hellermann and Vergun (2007) suggest, language samples from daily conversations of native speakers can also be used to highlight the use of discourse markers by fluent speakers or giving students adequate time for pair and small-group interaction in class can also foster the natural use of discourse marker for appropriate contexts (see Appendix E2). Sample activities to improve the use of discourse markers in classroom settings are suggested within the study (see Appendix E).

Moreover, non-native teachers of English also play a significant role in making learners aware of discourse markers and their pragmatic functions. Their use of discourse markers during courses and activities can be a model for the students. Thus, nonnative teachers should also be competent in the use of these pragmatic markers. As the study was conducted to the students who are teacher trainees, they will be English teachers of numerous language learners. Thus, they should be competent in discourse markers so as to make their students aware of them.

Another implication is about the course materials; specifically, textbooks. Course materials should provide examples of how target language is taught and how language is actually used in that specific community. Lam (2009) found out that there are wide discrepancies between teaching materials and naturally-occurring examples. Thus, course materials should be assessed to what extent they reflect authentic use of discourse markers.

In conclusion, it is suggested to increase language learners' awareness towards the variety and the use of discourse marker. Discourse markers are significant in making learners feel secure in speaking, enhancing in fluency, having naturalistic skills and becoming more native-like in spoken English. Thus, students should be exposed to more natural contexts with variety of functions of discourse markers and spontaneous conversations with pair or group discussions. Increased awareness and classroom practice on discourse markers will gradually help learners have discourse-pragmatic competence in English.

5.4. Further Research

Given the implications of the study, the study brings forth several suggestions for further research. Discourse markers as a research topic is a vast and prolific one in which plenty of research can be done. Although this study does not imply an overall generalization for any non-native speaker community, it has revealed the need to teach the use of discourse markers in English language teaching. Many studies can be conducted along this line, which will have considerable contribution to the development of pragmatic competence of language learners.

As the study has aimed to identifying the discourse markers in Turkish nonnative spoken discourse, a study which includes informed sessions about the use of discourse markers in language classrooms to observe the progress of the frequency, variety and functional use of discourse markers by nonnative speakers can be done.

Moreover, the use of discourse markers by nonnative teachers of English is as important as the use of discourse markers by nonnative students of English. Thus, a study can be conducted about nonnative teachers of English to find out whether they use discourse markers in instructional settings adequately and functionally or whether they are aware of the significance of these pragmatic markers.

Since this study is limited only to the discourse markers used in student presentations, future studies could use different discourse settings. For example, the use of discourse markers in more interactive atmospheres between students and teacher or among students such as lectures, group discussions or interactive tasks can be searched.

Another study can be done about the course materials to point out whether the materials or textbooks used in language classrooms are adequate to present learners with a variety of discourse markers and their pragmatic functions as course materials play an important role to help language learners reach authentic use of target language.

A final suggestion for further research is that studies can be done about the use of discourse markers by nonnative speakers whose native language is different from English. Thus, the use of discourse markers by nonnative speakers with different origins might have a contribution to comparative analysis among different groups of nonnative speakers of English and might be significant in the field of intercultural pragmatics.

REFERENCES

- Aijmer, K. (1987). Oh and ah in English conversation. In W. Meijs (ed), *Corpus Linguistics and Beyond*. (pp.61-86). Amsterdam: Rodopi.
- Aijmer, K. (2002). *English Discourse Particles*. Amsterdam: John Benjamins.
- Aijmer, K. & Anne-Marie, S-V. (2003). Well in English, Swedish and Dutch. *Linguistics*, 41(6), 1123-1161.
- Andersen, G. (1998). The pragmatic marker like from a relevance-theoretic perspective. In J. Andreas, H. Ziv, Yael (Eds.), *Discourse Markers: Descriptions and Theory* (pp.147-170). Amsterdam: John Benjamins.
- Andersen, G. (2001). *Pragmatic Markers of Sociolinguistic Variation: a Relevance-Theoretic Approach to the Language of Adolescents*. Amsterdam: John Benjamins.
- Anthony, L. (2005). AntConc: Design and development of a freeware corpus analysis toolkit for the technical writing classroom. *2005 IEEE International Professional Communication Conference Proceedings*. 729-37.
- Austin, J. L. (1962). *How to do things with words*. Cambridge: Harvard University Press.
- Bardovi-Harlig, K. (1996). Pragmatics and language teaching: Bridging pragmatics and language teaching. *Pragmatics and Language Learning*, 7, 21-39.
- Bardovi-Harlig, K. & Griffin, R. (2005). L2 pragmatic awareness: Evidence from the ESL classroom. *System*, 33, 401-415.

- Bazzanella, C. & Morra, L. (2000). Discourse markers and the indeterminacy of translation. In I. Korzen and C. Marengo (Eds), *On Linguistic Aspects of Translation* (pp.149-157). Alessandria: Edizioni dell'Orso.
- Bell, D. M. (2010). Nevertheless, still and yet: Concessive cancellative discourse markers. *Journal of Pragmatics*, 42(7), 1912-1927.
- Biber, D., Conrad, S. & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*. Cambridge: Cambridge University Press.
- Blakemore, D. (1987). *Semantic constraints on relevance*. Oxford, New York, NY: Blackwell.
- Blakemore, D. (1992). *Understanding utterances*. Oxford: Blackwell.
- Blakemore, D. (2002). *Relevance and linguistic meaning: The semantics and pragmatics of discourse markers*. Cambridge: Cambridge University Press.
- Blass, R. (1990). *Relevance relations in discourse*. Cambridge: Cambridge University Press.
- Du Bois, John W. (1991). Transcription design principles for spoken discourse research. *Pragmatics* 1, 71-106.
- Brinton, L. (1996). *Pragmatic markers in English. Grammaticalization and discourse functions*. Berlin/New York: Mouton de Gruyter.
- Brown, G. & Yule, G. (1983). *Discourse Analysis*. Cambridge: Cambridge University Press.
- Carter, R. & McCarthy, M. (1995). Grammar and the spoken language. *Applied Linguistics*, 16(2), 141-58.

- Castro, C. (2009). The use and functions of discourse markers in EFL Classroom interaction. *Profile 11*, 57-77.
- Celce-Murcia, M. & Olshtain, E. (2000). *Discourse and context in language teaching: A guide for language teachers*. Cambridge: Cambridge University Press.
- Chaudron, C. & Richards, J. (1986). The effect of discourse markers on the comprehension of lectures. *Applied Linguistics*, 7(2), 113-127.
- Chen Y. & A.W. He. (2001). Dui bu dui as a pragmatic marker: Evidence from Chinese classroom discourse. *Journal of Pragmatics*, 33, 1441-465.
- Clark, H.H. & Fox Tree, J. E. (2002). Using uh and um in spontaneous speaking. *Cognition*, 84(1), 73-111.
- Coll, M. U. (2009). *'Anyway': A formal Approach to the Syntax and Semantics of Discourse Markers*. Unpublished PhD Dissertation, Department of Language and Linguistics, University of Essex.
- Craig, R. T., & Sanusi, A. L. (2000). "I'm just saying": Discourse markers of standpoint continuity. *Argumentation*, 14(4), 425-445.
- Crozet, C. (2003). A conceptual framework to help teachers identify where culture is located in language use. In J.Lo. Bianco and C.Crozet (Eds.), *Teaching invisible culture: Classroom practice and theory*. (pp.34-39) Melbourne: Language Australia.
- Crystal, D. (1992). *An Encyclopedic Dictionary of Language and Languages*. Oxford: Blackwell.
- Cuenca, M-J. (2008). Pragmatic markers in contrast: The case of well. *Journal of Pragmatics*, 40 (8), 1373-1391.

- Curran, C. A. (1976). *Counseling-Learning in second languages*. Apple River: Apple River Press.
- Erman, B. (1987). *Pragmatic expressions in English. A study of you know, you see and i mean in face-to-face conversation*. Stockholm: Almqvist & Wiksell.
- Finocchiaro, M. & Brumfit, C. (1983). *The functional-notional approach: From theory to practice*. New York: Oxford University Press.
- Foster, P., Tonkyn, A., & Wigglesworth, G. (2000). Measuring spoken language: a unit for all reasons. *Applied Linguistics*, 21(3), 354-375.
- Fortuno, B. B. (2006). *Discourse markers within the university lecture genre: A contrastive study between Spanish and North-American lectures*. Unpublished PhD Dissertation. Universitat Jaume I, Castellon.
- Fox Tree, J. E. & Schrock, J. C. (2002). Basic meanings of you know and I mean. *Journal of Pragmatics*, 34(6), 727-747.
- Fraser, B. (1988). Types of English discourse markers. *Acta Linguistica Hungarica*, 38(1-4), 19-33.
- Fraser, B. (1990). An approach to discourse markers in English. *Journal of Pragmatics*, 14(3), 383-395.
- Fraser, B. (1999). What are discourse markers?. *Journal of Pragmatics*, 31(7), 931-932.
- Fuller, J. M. (2003). The influence of speaker roles on discourse marker use. *Journal of Pragmatics*, 35(1), 23-45.
- Fung, L. & Carter, R. (2007). Discourse markers and spoken English: native and learner use in pedagogic settings. *Applied Linguistics*, 28(3), 410-439.

- Gattegno, C. (1972). *Teaching foreign language in schools: The silent way*. (Second Edition). New York: Educational Solutions.
- Giora, R. (1997). Discourse coherence and theory of relevance: stumbling blocks in search of a unified theory. *Journal of Pragmatics*, 27(1), 17-34.
- Giora, R. (1998). Discourse coherence is an independent notion: a reply to Deirdre Wilson. *Journal of Pragmatics*, 29(1), 75-86
- Grice, H. P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds), *Syntax and Semantics*, (pp. 113-127). Volume 9. New York: Academic Press.
- Grice, H. P. (1989). *Studies in the way of words*. Cambridge, MA: Harvard University Press.
- Gupta, A. F. (1995). 'Singaporean English: Pragmatic particles. *Speech and Hearing Association (Singapore) Newsletter*. Retrieved on 1 June 2010 from www.leeds.ac.uk/english/staff/afg/shastxt.html/#Three.
- Halliday, M.A. K. & Hasan, H. (1976). *Cohesion in English*. London: Longman.
- Hansen, M. M. (1997). *Alors* and *donc* in spoken French: A reanalysis. *Journal of Pragmatics*, 28, 153-187.
- Hansen, M. M. (1998). The semantic status of discourse markers. *Lingua*, 104(3-4), 235-260.
- Hasselgren, A. (2002). Learner corpora and language testing. Smallwords as markers of learner fluency. In S. Granger, J. Hung, and S. Petch-Tyson (Eds.), *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching* (pp. 143-173). Amsterdam: John Benjamins.

- Hays P.R. (1992). Discourse markers and L2 acquisition. In D. Staub and D.Cheryl. (Eds.), *The Proceedings of the Twelfth Second Language Research Forum*. Michigan: Papers in Applied Linguistics. 24-34.
- Hellermann, J. & Vergun, A. (2007). Language which is not taught: The discourse marker use of beginning adult learners of English. *Journal of Pragmatics*, 39(1), 157-179.
- Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge: Cambridge University Press.
- Hussein, M. M. (2009). *Relevance theory and procedural meaning: the semantics and pragmatics of discourse markers in English and Arabic*. Unpublished PhD Thesis. School of English Literature, Language and Linguistics, Newcastle University.
- Iten, C. (1998). The meaning of although: a relevance theoretic account. *UCL Working Papers in Linguistics*, 10, 81-108.
- Johnson, F. & Paulston, C.B. (1976). *Individualizing in the language classroom*. Cambridge: Mass. :Jacaranda.
- Jucker, A. H. (1993). The discourse marker well: A relevance-theoretical account. *Journal of Pragmatics*, 19(5). 435-452.
- Jucker, A. H. & Smith, S. W. (1998). And people just you know like ‘wow’: Discourse marker as negotiating strategies. In A. H. Jucker and Y. Ziv (Eds.), *Discourse Markers: descriptions and theory*. (pp.171-201). Amsterdam: John Benjamins Publishing Company.
- Jucker, A. H. & Y. Ziv. (1998). Discourse markers. Introduction. In A. H. Jucker & Y. Ziv (Eds.), *Discourse Markers: Descriptions and Theory*. Amsterdam: John Benjamins.

- Knott, A. & Dale, R. (1994). Using linguistic phenomena to motivate a set of coherence relations. *Discourse Processes*, 18(1). 35-62.
- Krashen, S. D. (1981). *Second language acquisition and second language learning*. Oxford: Pergamon.
- Krieger, D. (2003). Corpus linguistics: What it is and how it can be applied to teaching. The Internet TESL Journal, 9(3). Retrieved on 25 June 2010 from www.iteslj.org/Articles/Krieger-Corpus.html .
- Lam, P. W. Y. (2009). Discourse particles in corpus data and textbooks: the case of well. *Applied Linguistics*, 31(2), 260-281.
- Lee, I. (1999). The acquisition of colloquial features by Korean Americans. *Journal of the Pan-Pacific Association of Applied Linguistics*, 3, 71–87.
- Lee-Goldman, R. (2010). No as a discourse marker. *Journal of Pragmatics*, 43(10). 2627-2649.
- Leech, G. (1991). The state of the art in corpus linguistics. In K. Aijmer and B. Altenberg (Eds.), *English Corpus Linguistics: Studies in Honour of Jan Svartvik*, (pp.8-29). Longman: New York.
- Lenk, U. (1997). Discourse markers. In J. Verschueren, J-O. Östman, J. Blommaert & C. Bulcaen (Eds.), *Handbook of Pragmatics* (pp.1-17). Amsterdam: John Benjamins.
- Levinson, S. C. (1983). *Pragmatics*. Cambridge: Cambridge University Press.
- Liao, S. (2009). Variation in the use of discourse markers by Chinese teaching assistant in the US. *Journal of Pragmatics*, 41(7), 1313-1328.
- Li-juan, Q. (2007). Grice's theory of conversational implicature. *Sino-US English Teaching*. 4(7).

- Louwerse, M.M. & Mitchell, H.H. (2003). Towards a taxonomy of a set of discourse markers in dialog: a theoretical and computational linguistic account. *Discourse Processes*, 35, 199-239.
- Maschler, Y. (1994). Metalanguaging and discourse markers in bilingual conversation. *Language in Society*, 23(3), 325–66.
- Maschler, Y. (1998). Rotse' lishmoa ke'ta? 'Wanna hear something weird/funny?' [lit. 'a segment']: Segmenting Israeli Hebrew talk-in-action. In A. H. Jucker & Y. Ziv (Eds.), *Discourse Markers: Descriptions and Theory* (pp.13-60). Amsterdam: John Benjamins.
- McCarthy, M. J. & Carter, R. A. (1995). Spoken grammar: What is it and how can we teach it?. *English Language Teaching Journal*, 49(3), 207-218.
- McCarthy, M. (1996). *Discourse analysis for language teachers*. (Sixth edition). Cambridge: Cambridge University Press.
- Müller, S. (2004). Well you know that type of person: functions of well in the speech of American and German students. *Journal of Pragmatics*, 36(1), 1157-82.
- Müller, S. (2005). *Discourse markers in native and non-native English Discourse*. Amsterdam/Philadelphia: John Benjamins.
- Nikula, T. (1993). The use of lexical certainty modifiers. In L. F. Bouton & Y. Kachru (Eds.), *Pragmatics and Language Learning, Vol. 4, 1993* (pp. 126–142). Urbana-Champaign: University of Illinois.
- Östman, J.-O. (1981). *You know: A discourse functional approach*. Amsterdam: John Benjamins.
- Quirk, R. (1959). *The teaching of English*. London: Secker & Worburg.

- Quirk, R., Sidney, G., Geoffrey, L. and Svartvik, J. (1985). *A comprehensive grammar of the English language*. London: Longman.
- Qun, Z. (2009). Fossilization of “so” as a discourse marker in cross-cultural communication. *US-China Foreign Language*, 7(8), 1-5.
- Redeker, G. (1990). Identical and pragmatic markers of discourse structure. *Journal of Pragmatics*, 14(3), 367-381.
- Redeker, G. (1991). Linguistic markers of discourse structure. *Linguistics* 29(6), 1139-72.
- Rose, K. R. & Kasper, G. (Eds.). (2001). *Pragmatics and Language Teaching*. Cambridge, NY: Cambridge University Press.
- Schiffrin, D. (1987). *Discourse Markers*. Cambridge: Cambridge University Press.
- Schiffrin, D., Tanen, D & Hamilton, H. E. (Eds.). (2003). *The handbook of discourse analysis*. (First edition). UK: Blackwell Publishing Ltd.
- Schourup, L. (1985). *Common discourse particles in English Conversation: Like, well, y’know*. New York: Garland.
- Schourup, L. (1999). Discourse Markers. *Lingua*, 107(3), 227-265.
- Searle, J. R. (1962). *Speech Acts: An essay in the philosophy of language*. Cambridge: Cambridge University Press.
- Silverman, D. (1993). *Interpreting qualitative data: Methods for analysing talk, text and interaction*. London: Sage.
- Siniajeva, I. (2005). *Discourse Markers: Their Functions and distribution across registers*. Unpublished MA Thesis. Vilnius Pedagogical University, Vilnius.

- SoundScriber. (1998). University of Michigan Regents. Michigan. URL: <http://www-personal.umich.edu/~ebreck/code/sscriber/>
- Svartvik, J. (1980). Well in conservation. In S. Greenbaum, G. Leech, & J. Svartvik (Eds.), *Studies in English Linguistics for Randolph Quirk* (pp.167-177). London: Longman.
- Stubbs, M. (1983). *Discourse analysis: The socio-linguistics analysis of natural language*. Oxford: Blackwell.
- Sun, Y. C. & Wang, L. Y. (2003). Concordancers in the EFL Classroom: Cognitive Approaches and Collocation Difficulty. *Computer Assisted Language Learning*, 16 (1), 83-94.
- Terrell, T. P. (1977). A natural approach to the acquisition and learning of a language. *Modern Language Journal*, 61(7): 325-36.
- Trujillo Sáez, F. (2003). Culture in Writing: Discourse Markers in English and Spanish Student Writing. In *Tadea seu liber de Amicitia*, Departamento de Didáctica de la Lengua y la Literatura (Eds.), (pp 345-364), Granada::Imprenta Generalife.
- Taboada, M. (2006). Discourse markers as signals (or not) of rhetorical relations. *Journal of Pragmatics*, 38(4), 567-592.
- Taylor, C. (2008). What is corpus linguistics? What the data says. *ICAME Journal*, 32.
- The MICASE (Michigan Corpus of Academic Spoken English). (2001). R. C. Simpson, S. L. Briggs, J. Ovens & J. M. Swales. URL: <http://micase.umdl.umich.edu/m/micase/>
- Thomas, J. (1983). Cross-cultural pragmatic failure. *Applied Linguistics*, 4(2), 91-112.
- Trillo, J. R. (2002). The pragmatic fossilization of discourse markers in non-native speakers of English. *Journal of Pragmatics*, 34(6), 769-784.

- Weinert, R. (1998). Discourse organisation in the spoken language of L2 learners of German. *Linguistische Berichte*, 176, 459-488.
- Wichmann, A. & Chanet, C. (2009). Discourse markers: a challenge for linguists and teachers. *Nouveaux cahiers de linguistique française*, 29, 23-40.
- Wierzbicka, A. (1991). *Cross-cultural Pragmatics. The Semantics of Human Interaction*. Berlin: Mouton de Gruyter.
- Wilkins, D. A. (1976). *Notional syllabuses: A Taxonomy and its relevance to foreign language curriculum development*. Oxford: Oxford University Press.
- Wilson, D. & Sperber, D. (2004). Relevance theory. In L.R. Horn & G. Ward, (Eds.). *The Handbook of Pragmatics* (pp. 607-532). Oxford: Blackwell. Retrieved on 11 March 2011 from <http://www.dan.sperber.fr/?p=93>.
- Wilson, D. & Sperber, D. (1995). *Relevance: Communication and cognition*. (Second edition). Oxford: Blackwell.
- Wong, J. (2000). The token “yeah” in nonnative speaker English conversation. *Research on Language and Social Interaction*, 33(1), 39-67.
- Zwicky, A. (1985). Clitics and particles. *Language*, 61(2), 283-305.

This image shows a full page of white paper with horizontal dotted lines, typical of primary school writing paper. The lines are evenly spaced and run across the entire width of the page. There are no margins, text, or other markings present.

APPENDIX B: MICASE TRANSCRIPTION AND MARK-UP CONVENTIONS

SGML TAG or SYMBOL	MEANING/ DESCRIPTION	APPEARANCE IN ON-LINE TRANSCRIPTS (HTML VERSION)
SPEAKER ID		
<U WHO=S1>, <U WHO=S2>, etc.	Speaker IDs, assigned in the order they first speak.	S1: at the beginning of each turn or interruption/backchannel.
<U WHO=SU>, <U WHO=SU-f>, <U WHO=SU-m>	Unknown speaker, without and with gender identified	SU: SU-f, SU-m
<U WHO=SU-1>	Probable but not definite identity of speaker	SU-1:
<SS>	Two or more speakers, in unison (used mostly for laughter)	SS:
PAUSES		
<PAUSE DUR=:05>	Pauses of 4 seconds or longer are timed to the nearest second.	<P: 05>
,	Comma indicates a brief (1-2 second) mid- utterance pause with non-phrase- final intonation contour.	,
.	Period indicates a brief pause accompanied by an utterance final (falling) intonation contour; not used in a syntactic sense to indicate complete sentences.	.
...	Ellipses indicate a	...

pause of 2-3 seconds		
OVERLAPS		
<OVERLAP>XXX</OVERLAP>	This tag encloses speech that is spoken simultaneously, either at the ends and beginnings of turns, or as interruptions or backchannel cues in the middle of one speaker's turn. All overlaps are approximate and shown to the nearest word; a word is generally not split by an overlap tag.	Text of overlapping speech is in blue.
BACKCHANNEL CUES and FAILED INTERRUPTIONS		
Embedded utterance (<u>tag within a <u> tag)	Backchannel cues from a speaker who doesn't hold the floor and unsuccessful attempts to take the floor are embedded within the current speaker's turn, and not shown as a separate line/paragraph.	[S3: Text of embedded speech is in orange and surrounded by orange square brackets.]
Embedded and overlapped utterance (<OVERLAP> tag within an embedded utterance)	Backchannel cues or unsuccessful interruptions that overlap with the main speaker's speech.	[S3: Text of embedded speech that is overlapped is in blue and surrounded by orange speaker ID and square brackets.]
LAUGHTER		
<EVENT DESC=LAUGH> or <EVENT DESC=LAUGH WHO=S2>	All laughter is marked. Speaker ID not marked if current speaker laughs.	<LAUGH>, <S8 LAUGH>, <SS LAUGH>, etc.

CONTEXTUAL EVENTS		
<EVENT DESC="WRITING ON BOARD">	Various contextual (non-speech) events are noted, usually only when they affect comprehension of the surrounding discourse.	<WRITING ON BOARD>
<EVENT DESC="APPLAUSE">		<APPLAUSE>
<EVENT DESC="AUDIO DISTURBANCE">, <EVENT DESC="BACKGROUND NOISE">		<AUDIO DISTURBANCE>, <BACKGROUND NOISE>
<EVENT DESC="SOUND EFFECT">, <EVENT DESC="GASP">		<SOUND EFFECT>, <GASP>
READING PASSAGES		
<SEG TYPE="READING">XXX</SEG>	Used when part of an utterance is read verbatim.	<READING>XXX</READING>
FOREIGN WORDS		
<FOREIGN>XXX</FOREIGN>	Used for non-English words or phrases.	Italics e.g.: the mother says c'est quoi? and Annika says to parce que eh and then,...
PRONUNCIATION VARIATIONS		
<SEG TYPE="PRON" SUBTYPE="/seltik/">Celtic </SEG> >	Used when an unexpected pronunciation is used that would affect comprehension of the surrounding discourse. Dialect or other phonological variations are generally not represented.	Pronunciation guide follows the word e.g.: ...they asked the librarian for pictures of old Celtic <PRON: /seltik/> uniforms the basketball team, and it turns out that the project was he was supposed to find Celtic <PRON: /keltik/> costumes.
<SIC>XXX</SIC>	Used when a speaker makes a mistake without self-correcting, and the error might otherwise appear to be a transcribing error.	(sic) follows the word. e.g.: despite the fact that that was the era of Women's Liberation like i say on the cover of Newsweek, and Gloria Steinman (sic) and uh Betty Friedan...

UNCERTAIN or UNINTELLIGIBLE SPEECH		
(xx)	Two x's in	i don't (xx) whole (xx) analysis it
(words)	parentheses	just struck me...
	indicate one or	lemme not write it that way (lest it
	more words that	be confused) with C syntax...
	are completely	
	unintelligible.	
	Words	
	surrounded by	
	parentheses	
	indicate the	
	transcription is	
	uncertain.	

(taken from <http://micase.elicorpora.info/micase-statistics-and-transcription-conventions/micase-transcription-and-mark-up-convent>)

APPENDIX C: SAMPLE TRANSCRIPTS FROM RESEARCH CORPUS

C.1.

Title: Sociolinguistics: Social Class

Gender of the presenter: Female

Recording Date: April 22, 2011

Recording Duration: 15 minutes

Word Count: 1.354

S1: yes, yes?

S2: okay, today i'll mention about social class. uhh the social class term is used for the differences between people which are associated with differences in social prestige, wealth and education. uhh for example, lawyers don't speak in the same way as the burglars they defend uhh class divisions are based on such status differences, status, status refers to difference or respect people give someone or don't give them, uhh status derives in western society from the material resources uhh a person can comment. also family background may be a source of status indepen- independently of wealth, for example the youngest child of and (xx) may be poor but respected, uhh class is used for groups of people who share similarities in uhh economic and social status, hmm people from different uhh social classes speak differently uhh the most obvious differences stare in vocabulary, in ninety uhh fifteen fift- fifties in England, many pairs of words were identified which it was claimed distinguished the speech of upper class english people u-speakers upper classes u-speakers from the rest non-u speakers if these vocabulary differences (xx) at all they are rather like those which distinguished Brahmin and non-Brahmin castes, uhh they disting- distinguished social group on a categorical basis. uhh for example uhh u uhh words and non-u words, for example by per by bicycle (xx) jacket uhh dress suit new uhh jack vegetables, green, ice, ice cream, sand, perfume, perfume (xx), okay the others. [S2: mental] men-mental.

<SS: UNINTELLIGIBLE SPEECH>

S3: the difference.

S4: could have tears. <LAUGHTER>

S5: okay could have (xx).

S2: upper class and non-upper class.

S3: hmm. okay <UNIDENTIFIED SPEECH>

S2: okay. please uhh, uhh read the book umm. example ten. look at example ten, you can see these non uhh u-words in it. please read the example ten.

<PAUSE: 36>

<SS: UNINTELLIGIBLE SPEECH>

S2: yes.

<SS: UNINTELLIGIBLE SPEECH>

<PAUSE: 9>

S3: we did.

S2: have you finished?

S4: yep. S5: yes.

S2: okay. umm. okay which words umm do you find?

S3: note paper.

S2: example for non-upper class. note paper, yes. for example handbag.<SS: UNINTELLIGIBLE SPEECH> settle. okay.

S4: perfume.

S5: yes perfume... and the barriers uhh barriers between groups are not in term mountable as in castled-based soc- uhh societies people can move up or down the social ladder and this potenti-potential modelity is mirrored more accurately in other aspects of their speech such as pronunciation and pronunciation, groups are often uhh distinguished by the frequency with which they use particular features...rather than uhh by their use of completely different forms ...

S6: *çok pardon ...*

S2: please look at example uhh twelve uhh twelve uhh (xx) radio broadcast of two elderly peoples (xx) childhood in Edward Dean Ballington. the two speakers contrast on a range of social variables. please read the example twelve.

<PAUSE: 33>

S2: have you read it?

S7: yes.

S2: okay. hmm. the two recordings were analyzed to see whether there were uhh differences between the speakers in terms of the number of hs uhh they dropped. ucha found that (xx) did not omit a single h while george davis of eighty three percent of the hs which occured in his interview. Uhh (xx) is in uhh upper class and uhh george davis in is in low social class uhh so uhh (xx) did not omit a single uhh h while George Davis drop in drop eighty six percents of the hs. the speakers' different social backgrounds were cl- clearly reflected in this feature of their speech. this speech variable is variedly called h-dropping... uhh

figure six point three shows the average h-dropping scores for five different social groups in two different places in England, West Yorks- Yorkshire and Norwich uhh the first group is uhh upper class and the uhh group five is uhh low social class for wor- hmm wor- working class. uhh as you see umm in (xx) in both areas the highest social group one uhh drops the least number of hs uhh and the lowest group uhh five on (xx) most.

S7: yes.

S2: okay. there are regional differences in that the west yorkshire's scores are systemat- systematically higher than the Norwich's scores. but the overall pattern rema- reminds the same. uhh Norwich and there are regional differences but the umm average uhh result is same. uhh in the West Yorkshi- Yorkshire study for example one person umm. sorry, also these are averages and within each social group there is always (xx) of individual variation. in the West Yorkshire study for example one person who belon- belong social in the middle group three uhh drop every h. William Lavolt interviewed uhh one hundred twenty people and examined their pronunciation of a number of different consonants and vowels. he found regular patterns related to the social class of the speakers to the percentage of standard as opposed to vernacular pronunciation they produce. for instance the pronunciation ing uhh and in at the end of verse like uhh sleeping and swimming distinguishes social groups in every in this speaking community in which it had been investigated... table six point two percentage of of vernacular in pronunciation for four social groups in speech community bright Britain, America and Australia uhh again first group is uhh upper class and four group is umm low social class, working class. uhh in each community people from lower social groups use more of the vernacular in variant than those from higher groups. as with h-dropping there are regional var- regional variations might be. uhh between communities but the regularity of the sociolinguistic pattern in all for communities is quite clear, as you see.

SS: hmm...

S2: r pronunciation... there are two possible variation of r uhh either it is present and pronounced r or it's absent. in some regions pronouncing r is part of the standard prestige dialect. for example in Scotland, in Ireland, in the Bo- Boston and Newyork in other areas standard dialect speakers don't pronounce r after vowels. in areas where r pronunciation is prestigious uhh the higher persons social group group, the more r they pronounce. the higher uhh a person's social group the more they pronounce. uhh please look at uhh example thirteen.

<PAUSE: 42>

S2: have you read it?

SS: hh1h. ..

S2: okay, umm (xx) went went to the three Newyork city department stores uhh that provided to three (xx) socio-economic groups uhh (xx) avenue and expensive upper middle class store. mention is a less expensive middle class store (xx) is a discount store frequently mainly by working class Newyorkers. he studied how they pronounce the uhh phrase (xx) then pretend pretending he hadn't heard the answer, he said excuse me uhh people repeated the their answers and he obtained the second and more careful pronunciation uhh here is the result demon- demonstrated that the employees at Susse high social group used post-vocalic r more often, as you see. more often. Mershies middle social class employees used a less often, less often than uhh Susse upper social class. claim uhh low social class employees rarely used r... okay. post-vocalic r ill- illustrates very clearly the arbitrariness of the particular phone which are considered standard of presitigious while pronouncing r is considered prestigious in Newyork city in a region in England is, it is not. uhh uhh you can look at uhh six point four...table

six point four. these reflected in the pattern for the different social groups in two cities illustrated in figure six point four. in Newyork city... uhh the higher your social class the more you pronounce post-vocalic r. in the reading the higher your social class, the fewer you pronounce post-vocalic r. as i mean it can be changeable... uhh the arbitrariness is illustrated by h-dropping uhh Kim's comment in please uhh look at example eleven, page one hundred thirty six. example eleven. seven sorry..

<PAUSE: 19>

S2: have you read it?

SS: hmm.

S2: okay, Kim's comment in example uhh eleven expressive by the held viewpoint on a only educated people drop their hs. but in the center the top social class in England in England drop the h at the beginning of words. clearly the particle linguistic forms which people regard as prestigious or stigmatized are in general totally, totally arbitrary but most probably uhh the speech of the most prestigious social group determine the uhh price values of standard forms... that's all... are there any question?

SS: no.

C.2.**Title:** Social class: Vowels.**Gender of the presenter:** Female**Recording Date:** April 22, 2011**Recording Duration:** 11 minutes**Word Count:** 1.294

S1: okay, as you see here, today i'm going to talking about, i'm, i'm going to talk about vowels, other languages and grammatical patterns according to social class. okay. firstly i, am, i want to start with a, an example. please read it. <PAUSE: 15> have you read? [S2: no. S3: one minute.] no? okay.

S3: one minute.

<PAUSE: 16>

S1: okay?

SS: yes.

S1: what do you think the writer's complaining is about, about?

S3: umm. the usage of i. some vowel.

S4: yes. i am.

S5: the letter i is changed into you. [SS: you] yes.

S1: yes. you meant the way pronounce their vowels. uhh as you see here, they are New Zealanders. and uhh they pronounce till as tall and him because hum so the writer is complaining about their pronouncing in vowels. uhh measuring them, measuring small differences in the way speakers pronounce the same vowels is very challenging task. uhh and because of this, (Lavolt) developed a method to measure it by giving a score to different

pronunciations according to how close uhh they were to the prestige pronunciation or standard in the community. there are, there is a scale and there are four peak points on it. uhh and they were used to measure different pronunciations, a score four wa- uhh was for the pronunciations close closes to rp. what is rp, by the way?

<SS: UNINTELLIGIBLE SPEECH>

S1: yes, the false show class accent and uhh point fi- point one to the broadest New Zealand pronunciations with two bi- with two points in between. uhh New Zealanders consider rp in an appropriate standard accent for them but uhh in practice it is still an (xx) prestige norm uhh they deny it but.

S6: they use it.

S1: they use it. the scores reveal the social basis of New New Zealand patterns of pronunciation. i mean, the higher a person social class the closer their pronunciation was to rp. and other languages, although the social linguistic patterns have been mostly researched the English speaking communications commun- communities they have been founded other languages. for example in Paris, they pronounce the pronunciation of the first vowel in uhh in verse varies from one social group to another. in Montreal, the frequency with which l is deleted distinguishes the French of two social groups as in the [S7: table]. table. for example uhh *ill* for impersonal it means it in English. uhh in profession in professional cla-class the usage is uhh lower and working class the usage is higher. and personal in uhh for personal it means he in professional class uhh again the usages lower from the working classes' uhh usage. and *elle* again the *ce* it is lower than the working class usage. it shows that not only does *elle* deletion differ between the social classes. it also differs according to the grammatical status of the word in which it occurs. *elle* almostly dissapeared in Montreal French in pe- in impersonal *elle* in. i mean, it. and the surrounding sounds

also affect the life (xx) elle deletion. it is much more like to the disappear before a consonant then, before a vowel. most people drop their hs in an un- unstressed level. so, uhh sorry here i want you to look at example eleven on page one hundred and thirty seven. <PAUSE: 6> it's an example of deleting deleting hs.

S8: example twenty?

S1: no, eleven.

S8: eleven.

<PAUSE: 16>

S1: so deletion of h is a symbol of uhh that you are a.

<SU: UNIDENTIFIED SPEECH>

S1: uhh yes, yes. uhh so linguistic as well as social factors are relevant in accounting, accounting for patterns of pronunciation. uhh similar patterns can be found in any speech community where there are social classes, in Tahrani Persian as well uhh as well as in which (xx) use in Mombasa the same relationship in, is found between speech and social class. the higher, again the same thing, the higher social groups use more of the standard forms while the lowest group use the fewest standard forms. grammatical patterns again we have a, we have an example. uhh there is a girl in a_ is_, she is eighteens she's getting story effect here she has seen (xx) again. there are some mistakes in this example...

S8: please (xx)

S1: please read it yes. please read it.

<PAUSE: 16>

S1: have you read?

SS: yes.

S1: what is the mistake?

S3: for example play and play.

S1: h1-h1h. yes.

[**S4:** theirs.]

[**S1:** and umm sorry.]

SS: theirs.

S1: all stare..was making umm.

S5: then the little flies was making.

S1: yes. so, uhh on average it was found that children from lower class the Malaysed used more vernacular word forms than children from middle class families. this pattern has been noted for a variety of gram- grammatical variables. again here are some examples of standard and vernacular uhh grammatical forms. for example, in past and word forms <READING> "i finished that book yesterday". it is the correct one. [**S8:** h1-h1h]. right? and second form. "i finished that book yesterday."

S6: it's false.

S1: yes, it's false. and present tense word forms <READING> "Rose walks to the school everyday, Rose walk to school everyday". that omit as here. and negative forms "nobody wants any chips, no body don't want no any chips". what is the mistake here?

<SS: UNIDENTIFIED SPEECH>

S1: actually we don't uhh use two negative.

SS: yes.

S1: found in this examples. and eight, "Jim isn't stupid". and "Jim ain't stupid". uhh as with pronunciation, there is clear pattern to the relationship between the grammatical speech forms and the social groups who use them. uhh and on our books, there is a diagram uhh please look at the, sorry, page.

SU: one hundred and thirty one.

S1: page one hund- yes one hundred and forty five. [SS: five]. figure six point five. it is diagram of this sentence. "those walked to school every day, those walk to school every day". omitting s. [SS: s] as you see, Aylin hocam, please <LAUGHS>.

<SS: LAUGHTER>

S1: as you see, in the diagram, uhh in the social umm highest, higher social class, there is no usage of the vernacular verbs uhh i mean they don't use she walk. they use, she walks and the higher social groups use more of the standard grammatical form and fewer instances of the vernacular or non-standard, non-standard form uhh where standard English allows only one negative in each class, as we do normally, most vernacular dialects can have two or more, as you, see in this sentence six, "nobody don't want no chips". we use which one, nobody wants any chips.

S8: only one negation.

S1: these are vernacular form and the other standard form, uhh and in some dialects, every possible form which can be negated

is, negated. and sorry. "a an adolescent gang member in New York produced the following. it ain't no cat (xx)". it means in standard English there isn't any cat bad then getting to any cook. and uhh an adolescent in Detroit "we ain't had no trouble about (xx) pulling out no knives". as you see here, they are what? they are, vernacular forms.

SS: yes, mutliple negation.

S1: and umm, they are multiple kind of they are multiple negation in the sentence. multiple negation is a grammatical construction which has been found in all English speaking communities, where a social dialect study has been done. umm in every community study, it is much more frequent in lower class speech top in middle class speech. and umm multiple negaton is a very salient. i mean clear, vernacular form. people notice it when it is said, even most. uhh middle class speakers tend to avoid it as we said before. while lower class speakers use it more comfortably. these usages of multiple negation replaces the salience. umm many factors interact in the term (xx) preparation of vernacular of standard forms a person uses. some of these are social factors such as age and gender and, and uhh some of them is the linguistic (xx) in which avert occurs and that's all.

S9: thank you.

C.3.**Title:** Gender and Age.**Gender of the presenter:** Female**Recording Date:** April 20, 2011**Recording Duration:** 11 minutes**Word Count:** 1.279

S1: okay, uhh today i'm going to talk about the difference in speech according to the gender and age. uhh, i want to start (xx) with a question. do women and men from the same speech community use different linguistic forms?

SS: yes. yes.

S1: yes. how? h1h. let's learn. firstly, i want to say the differences between sex and gender. sex refers with categories distinguished by biological characteristics. and gender is more appropriate for distinguishing people on the basis of their socio-cultural behaviour including speech so we will use gender.. gender exclusives speech differences. it's in non-Western communities and there are two situations uhh first one is that women and men speak different languages in a commun- community. and the other is that uhh the language is shared by women and men but with some particular linguistic features. let's have a look. yes, h1-h1h. for the first situation, there is an example. for example, Tayona uhh is a, is an Ama- Amazonian Indian woman. her first language is Desano. the language of her long house is Tuyuca. and the language of all men in this tribe uhh uses Tuyuca. and language she uses to talk to her children is again Tuyuca. but the language she uses to talk her husband is Desano. and the language her husband uses to talk to her is Tuyuca.

SU: hah!

S1: it's different. and the second situation, the language is shared by women and men but with some particular linguistic

features, uhh particular linguistic features occur only in the women's speech or only in the men's speech. these linguistic features occur only in the small differences in firstly pronunciation. for example, in Montana, an Indian tribe, uhh men use *yatsa* but women uses *jabsa* for bread. and in Bengali, a language of India and man uses initial *n* but women use initial *l*... and in the word shape it diff- it differs men and women use different affixes. for example, in Yana some of the words used by men are longer because they use suffix. for example, women say *bah* and men *bana* for deer. women *ya* men *yana* per person. (xx) how to (xx) pronounce it. and this is (xx) or (xx) to take he might go away.

S2: yes, maam.

S1: yes.

S2: men use more words.

S1: yes. **[S2: on the contrary of Turkey.]** they need more words, hah. **[SS: LAUGHTER]** yes. and in vocabulary, men and women use different vocabulary items. for example in Japan, women use *otosa* and men use *oyashi* for father. **[SS: hmm.]** and women *onaka* men *hara* for stomach. women, *oyashi* men *umayi* for delicious and women uses use *taberov* men *ku* for each. they completely different. they use completely different vocabularies. and the last (xx) in pronounce some languages signal the gender of the speaker in the pronounce syst- system. in Japanese there are a number of words for *i*. varied in formality. women are generally more restricted to formal variants. for example, woman women say *atashi vatashi*. **[SS: LAUGHTER]** *vatakushi* the most formal one. but men *ore* and *bugu*. *bogu* <LAUGHS> the most casual one. so let's come to the conclusion. gender exclusives speech form reflects social status or power, power differences and gender exclusive social roles. for example, again, in Bengali society, a wife isn't permetably used her husband's name. she addresses him with the term *sancho*. it means do you hear?. when she refers

to him she uses circumlocution, circumlocution is to say something indirectly. again uhh Bengali wi- wife. her husband's name is Parah and it means to go from his star. she can call him *Tarah* so she uses the term *nokowotro*. it means heavenly body.

S3: *bizdeki* <UNIDENTIFIABLE SPEECH>

S1: yes.

<SS: LAUGHTER>

S1: it's like the. and gender preferential speech features uhh it's social dialect research. let's look at this example four. in Yorkshire England, the name of the teacher is Mrs. Hall she says. and the boys call the teacher Mis. Hall. girls call the teacher Miss Hall or Mis Hall. so in Western communities where women and men's social roles overlap, equal. the speech, speech forms they use also overlap. women and men do not use completely different forms, they use different quantities or frequencies of the same form... in all English speaking cities, women uses uhh more (xx) pronunciation, fewer (xx) pronunciation than men. i Montreal, French used by men and women is distinguished by the frequencies with which they pronounce l. both women and men delete l but men do so more often than women. so conclusion, women tend to use more standard forms than men do, while men use more of the vernacular forms than women do...sa so let's uhh look at the gender and social class example five uhh Linda's uncle is a plumber and he talks just like the other man on the building site where he works and he he says (hus) for house, (hom) for home (cok) and (bus). and Linda's aunt works in a shop uhh at home she talks a bit like her uncle for example she says music (xx) but she talks to customers just like a lady. house and home... linguistic features which differ in the speech of women and men in Western com- communities are usually features which also distinguished the speech of people from different social classes. in every social status, men use more vernacular forms than women. for example, in social interviews in Norwich, men

use more of the vernacular in form at the end of the words like (xx) and and (xx) than women. and this pattern was quite consistent accross five distinct social groups. let's look at this table. vernacular in by sex and social group in Norwich. in each group man uses more of the vernacular in form. here, man.

SS: yes.

S1: you see. and this is the highest gro- social group and the lowest social group. in the lowest and highest socia- highest social groups the women's speech is closer to the that of the men. uhh you can see for third and for (xx) social groups. and uhh in the same group than to the uhh to that of women in other groups. for example uhh this is women's speech and this is men's. it is more closer than to the woman. we can see. so class membership is more important than gender identity. however, this is not so true of women in group two. look at this. uhh this (xx) score for vernacular form is closer to that of the women in group one. yet this closer. than it is to that of men from their own group. okay. so, conclusion again, across all social groups women generally use more standard forms than men. as so corres- correspondingly men use more vernacular form than women. uhh for example, in that trait, multiple negation we have mentioned. for example, "i don't know nothing about it". uhh, it is more frequent in men's speech than women so it's vernacular. this is through in every social group but the difference is most prompt in the second highest group. in second highest group multiple negation score is thirty two percent to one percent for women. even in the lowest social group, men use a third more instances of multiple negation than women. ninety versus fifty nine percentage.. this wide widespread pattern is also evident from a very early age. for example, in several rural new England village, boys use more m and girls more ing forms. and studies in Boston and Detroit, boys used to more vernacular forms such as consonant cluster, simplification, last for last, tall through tall. in Edinburg, differences of this search were

observed in the pronunciation of girls and boys as young as six years old. umm. that's all.

APPENDIX D: SAMPLE TRANSCRIPTS FROM MICASE CORPUS

Title: Bilingualism Student Presentations

Transcript ID: STP355MG011

Academic Division: Humanities

Publisher: Michigan Corpus of Academic Spoken English, English Language Institute,
University of Michigan

Interactivity Rating: Mostly interactive

Number of Participants:

Students: 6

Speakers: 7

Recording Date: April 16, 1998

Recording Duration: 99 minutes

Word Count: 15.153

S1: okay, now basically this is uh, a study on um, i wanted to look at um, the bilingual population here at the University of Michigan, and i started out thinking i would just look at, whatever people i could find who were bilingual, and then i realized that it would be difficult to kind of, find anything out about it if i was just looking at whatever languages i came up with. so i ended up concentrating on two different languages, and um, i also concentrated on the easiest samples i could find, so i have a Spanish um, English bilingual sample, taken from a class actually of self identifying U-S Latinos. so that's something to consider, and then i have um, a sample from, a Hindi class, uh two Hindi classes actually here, which are mostly bilingual um heritage speakers of Hindi, or another Indian language. um, and, uh basically, no one other language is highly represented besides Hindi but i'll talk about the little that a little bit in a minute. um the things that i was really looking for were, uh, had to do with, theoretical affiliation, with um a language community, which might serve as a community using these uh measures, basically,

S2: Helen

S1: yeah?

S2: can you put the window, down?

S1: the which?

S2: the window, down?

S1: the window down? <NOISE DISRUPTION> see if i can get those um... <NOISE DISRUPTION><SS: LAUGH> um, that helps a little [SU-
f: (xx)] okay, um okay the first thing i was looking at is theoretical affiliation like do you think that you will be basically affiliated with this language community in this way in the future (that's what) (xx) um so first i ask, i just sort of_ these are just theoretical right though they have nothing to do with what we are doing now and i'll, talk a little bit about one other thing they ask that's about right now, in a minute but these're basically, the main things i was looking at. and i also looked at the importance of maintaining or increasing their proficiency in language A, um, for the vari- these various reasons. and, what i ended up doing in terms of figuring out what these meant and how they correlated was, looking at um, basically what i did for these scores or these answers, is i gave them two, points if they, said yes one point if they say maybe and zero points if they say no so i have an aggregate score for that, and i have um an average, score for this one. um, depending on various things and i also i'll show you a little bit about how those broke down, um... there's a lot of different things you can consider in terms of what will influence, you know how someone, perceives their own identity and how important it is for them to maintain their other language. um basically with students at the U-of-M, we're looking at people who, um, are basically social- socially mobile, so um, i sort of went in assuming that i would find people who would be more tending to be agents of language shift, rather than language retention, um in the sense that they, most of the students um, like in the in the Spanish, bilingual,

Spanish English bilingual sample, three out of eleven, had grown up in a Latino neighborhood or, i_ whatever, i i measured that just by saying, you know, do you h- did you have neighbors who were spoke Spanish did you have um classmates who spoke Spanish? and only three of them answered yes to either of those. so basically most of them did not grow up in a bilingual community, except for that of their parents, and i've done qualitative interviews i haven't, finished evaluating all of that yet but basically, for the qualitative, work it seems like most of the students, did indeed grow up in um an integrated community, and they often say well i spoke Spanish or Hindi with my parents friends or their childre- their par- my parent's friends' children, but i didn't, you know have friends at school who spoke it or whatever. um another thing to consider of course is the different language groups, um, Hindi obviously, uh, is not, as widespread in the United States as Spanish is, um so in terms of usefulness or, accessibility in the United States it's a little different. so i'm just gonna look at some differences between the two groups, for a moment here, um, this is the, answer to th- the second question which is about the importance of, um, th- these various reasons for maintaining or increasing, fluency in language A (alright,) um, and i've got this little, graph at the bottom, which basically tells you that, i- i've sort of labeled this so, um, basically for the Hindi class, seventy-five percent answered_ had an average of this or higher, the seventy-fifth percentile, half of them had this or higher, twenty-five percent, so you can see that the Hindi class has a much, um, lower range basically, of answers okay, and you can basically, see that that's from these two, numbers here, i circled them here, um, okay, um so for professional and academic reasons, Hindis, were rated not that much more- much lower than the Hindi speakers- than the Spanish speakers. um, and that was pretty much across the board, and, we can think of some obviously um reasons for that, uh... another interesting thing you can see there's some variability um i don't think that the rest of it's very significant though, um, the variability in the different ranges, um there is some variability though there is

still like, this is relatively low compared to the other scores, for Spanish speakers, to the academics but it's still relatively high, this was just- a rating of four just means important, anyway so, and then, just to look at the other, thing i was interested in measuring, this is the questions on theoretical affiliation, and i broke these down into proportions so this is like basically like thirty-six percent, (xx) fifty-five percent. um, so for this- for the Spanish um, the Spanish English bilingual sample, or the Spanish class i should just say um, basically, there's a lot more, um, higher ratings in the in the yes category like, a lot more people who said yes to um, to everything, and there's also very few people who said no to anything. you'll see here most people said at least maybe, you know, or yes. um and that again might have something to do with the also the sample is taken from a U-S Latino_ self-identifying U-S Latino class, and so students who are in this class presumably are already relatively invested in being a U-S Latino in the sense that they self-identify um, but at the same time, um the Hindi speakers, um, no one said no for example, no they would not marry a Hindi speaker, but um, you know two people i believe that represents two people who said they would never marry someone who was a Spanish English bilingual. basically. okay. um, and you can see again there's, a higher distribution for the Spanish speakers than the Hindi speakers for the aggregate score, like all these together, um, (let's) see and the other interesting things, th- only one person i think said yes they would live in a country where Hindi was spoken basically India right, um, whereas, um several people said yes they would, (live, in) live in a Spanish speaking country, definitely and um, let's see, the other interesting thing is that hardly anybody said that, they would not raise their children bilingually, (okay) in any of the samples there's two people total, in both samples that said that they, would would not raise their children bilingually. or maybe actually they didn't say would not they said maybe i won't. you know <LAUGH> and one of them like even wrote a comment in like i'm not sure if i'm proficient enough to bring my children up bilingually

<LAUGH> like i want to but, so basically you can see by just, looking at that that, basically for all of the students sampled they are all pretty invested in, thinking that bilingualism is a good thing, you know, over all, something they want to pass on to their children, in the interviews that comes out as well... um, so i was interested also in how these correlate i was sayi- thinking you know m- well maybe some people are interested in being part of this community the bilingual community, but they're not really interested in, um, having, uh, high, their (own,) high proficiency (in the) language, so there is a high, (um,) there is a high correlation between these two things so, this is the average, this is the average score of like the five point rating thing and this is the aggregate score of like yes maybe no, do you want to do this do you want to do that, um, and there's like a high correlation i think it's point, seven two

S3: sorry could you just be clearer on what's correlated with what

S4: Speaker information restricted

S1: okay um, this is the average interest so this is like, um rate, the following um, how important is it for you to maintain or increase language A and you can just put this up, uh, rate that on a scale of one through five basically for the following reasons. (xx) right there, that's the bottom one. (aggregate interest,) and then i just averaged their scores okay, and that's the scale , this is the aggregate score so they get two points for every yes and one point for every maybe, so somebody who has twelve answered yes to everything, okay, somebody who has a three probably only answered yes to one thing maybe, to another representative of that so, and there's, you know there's definitely variability you know there's, people who, it's really important to them to, to maintain language A and yet they have a pretty low theoretical affiliation but there is a correlation of um, uh, this is th- this- these are the questions (xx) um... i was interested in the marrying someone who speaks language A thing for the Hindi speakers and i actually ended up asking,

some people in the qualitative interview whether they were interested in arranged marriages, and i only had one person say they were intr- they were considering it still so, but um since that's, a big thing for recent immigrants (of course) especially... um okay... now another thing that, um, i realized when i started looking at this is i asked a question, are you involved in any religious organizations student organizations, classes, whatever in in- th- where people speak language A, and um just correlating with the religion thing, got actually a really pretty high, correlation, so this is just like yes or no d- do you participate in a religious organization (xx) (people) (xx) okay, um, and you see this is the average of the five so this is the, you know how important it is for you maintain the language, and you can see that, um, this is, these are both for the Spanish sample. so for the Spanish sample i mean these are pretty low numbers so it's kind of weird to see it in this form but this just basically means that, we are (xx) (black) (line) (notes) fifty percent, whatever answered higher than that, so you can see that everybody, who's yes involved in their religious organization they did all fives, for all of the treatments, so, um, and, here again you see that, like ha- half of, all this, like two of these students are all, gave almost all yesses basically for all those (added) questions um, this_ it's kind of weird to see this and these low numbers i tried to this like that

S3: have you got a misprint there? is one of them empty?

S1: no these are all Spanish classes these are, for religion or yes or no.

S3: aah (xx)

S1: and th- the in and this is oh this is the aggregate score for like, [S3: i see] do you imagine yourself imagine- marrying somebody who speaks language A. so this is just for the Spanish one, and they say a, a little different, uh, distribution with the, Hindi class, but, with similar overall results except for

the, um this is reversed so let me just do it one at a time, this is the aggregate score, so you see again there's a higher, (like) median, um, this is like for yes no maybe do you wanna marry somebody (you wanna,) do you- do you think you'll marry someone who speaks language A or whatever. um, that's definitely still higher for people who are involved in religious organizations now, but it's a little more spread out, (and i) (xx) the- the Spanish sample seemed pretty_ more polarized there's also fewer people in the Spanish sample who are involved in a religion, here there's, the opposite there's nine, uh students involved in a religion and four are not, so, but then these this is kind of, bizarre I don't know what to make of this but i'll just show it to you, um this is for the, uh, this is for the average grading of (reasons) and, there's this huge spread for the people who said yes, and then there's this really small spread of people who said no, and so that's- maybe this says that maybe religion doesn't correlate with um the importance of speaking Hindi as much for these students okay um, that's- i mean it's a small sample so it's hard to make these kind of conclusions but, it does seem like, it could be different than the Spanish sample where, there was a high correlation in both cases with being involved in religion. okay... so um basically, what i came up with i guess is that there is um, that language is really important, to these students in terms of their affiliation with this community bilingual community or their cultural community maybe in a a larger sense to say that because, the other thing about this is, they are taking this class and language which means that they are investing time and energy in learning this language but it also means that they don't_ they're not already super proficient in this language right, so they're not like, um, the ideal bilingual and we all know that that doesn't happen very often but, but they you know they're not like um, going from being already you know totally maybe a part of this community and then taking this class of course cuz they're wanting to improve it, um, especially for the Spanish speakers i would say that i guess just because when you talk about Hindi speakers i just imagine

it being more, i mean, if you're not even learning those character set in school you know, um and a lot of the_ from the qualitative interviews um a lot of what i got was that people, um rejected their their home language as a child, they'd often say oh i you know i told my parents to stop speaking Spanish to me you know all this stuff when they were a kid, um and but now i want to. you know so there's this kind of change in heart sort of coming with, age and um, the other thing is that, it's interesting to see why it is that students want to revive their other language in a sense, okay, and basically what i n- uh, what i decided is that they are basically interested in it because of um, they want to use it as a way to, be part of this um cultural identity these are (the same) quotes as i had before so people who've seen this already, it's boring but, um basically these are two interviews um where basically students said right out the reason i want to speak Spanish is because, i um, i want, i want to be considered Latino and i feel like this is, like, validates me in some way basically what they're saying so they, he says, because i don't look like a Latino person, since i'm not brown, uh let's see, are they wha- are your kids gonna be you know Hispanic, and so i wanna know you know i wanna speak Spanish basically. is kinda what it's coming down to and this student has the same basic idea she says, um, it's part of myself i really wanted to get back, that back, that back you know and then she says, basically i look African-American, and most of my friends have been Afro-American so i really haven't had that much, in, you know so she's basically getting this, idea that she wants to um be part of a Latino community, in some way, that she hasn't been so far and she feels like Spanish is gonna come- some how, help her, um her overall identity with that. so that's kind of what's going on with the qualitative data i haven't finished analyzing the rest of my qualitative data so that's all of, that for now. um, i'm trying to think, how am i doing on time?

S3: okay, we didn't really talk about timing how how much more have you got?

S1: how much more have i got? i, that's that's most of it i could say a couple more things but

S3: well

S1: okay

S3: do you want to briefly, a couple of minutes perhaps, [**S1:** okay] and i think we have twenty minutes for questions.

S1: get my

S3: i think we made it twenty minutes a piece, didn't we?

S1: yeah, i think and then time for questions, and then there's uh... (i don't) (have this) it's in my bag... um, the one other interesting thing is and i already talked about this in my other presentation but um there was this distribution lemme just find that, <LOOKS IN BAG> there was this distribution in terms of who asked people who they spoke Spanish or English with or who they spoke (xx) language A who do you speak language A with, and um... there was this thing that is typical of the bilingual population (at first) but it's interesting to see it, um, where you get the, speaking both with their parents fourteen out of this is for both language groups. fourteen out of uh, i guess twenty-four total spoke, speak both with their parents okay, six speak only language A with their parents, um, but with siblings you see a much higher rate of speaking only English okay so a lot of students told me this in interviews too they said i only speak English with my with my brother or whatever, um some speak both very few speak only language A. only one person only speaks language A with his friends uh i think that was a he, um, uh and eight only speak English, which is um, less in fact than people who only speak English with their um <S1: LAUGH> with their siblings but the thing about the friends thing i should just say that this- this whole friends line is kind of strange because, this um fifteen this the lines here this fifteen's kind of like we really don't know what that is it could be like i have one friend who speaks English and everyone else speaks language A,

or it could be like i have fifteen friends who speak English and one friend who speaks language A so that's that doesn't really tell us much but there are eight people who only have friends basically who speak English so that's that's a pretty large number. who are just not involved basically in a language A community at this point. in terms of their own peers um so. um, okay, my transparencies... <S1: FLIPPING THROUGH TRANSPARENCIES> okay. so basically my conclusions were that students who are more involved in a bilingual community or who see themselves becoming more involved in it in the future with these kind of imaginary questions, are are also more concerned with maintaining language A, be that Hindi or Spanish or Gujarati or there's um, i guess probably no one in here does Indian languages but i can mention that the different languages i was asked this before so the different languages that they spoke besides Hindi were Marathi Urdu two Gujarati and one Oriya. so there's kind of a wide range of other languages, and Arabic one other spoke Arabic. um, okay so basically there is this high correlation between seeing themselves being involved in this community and wanting to maintain language A. so that sort of says to me there's the language is an important part of their cultural identity and often students when i asked them that i said what is_ what d- does your culture mean to you um outside of the language you know they would kind of say well, there's food and you know there's, you know the religious customs they would often say but beyond that that was kind of it um, there were some other interesting things that came up. um, but in language A the the the language seemed to play a more important part in it. um let's see, and, let's see <PAUSE:07> and i'm sort of wondering and i don't know if i'll be able to answer that in this but i'm i'm sort of wondering what in what way the students hope to retain language A if they're not gonna be involved in the community and then you get into this more symbolic, orientation towards this community. where it's like i want to be whatever an- an Indian American, but um, maybe i'm just tend to spend more time with you know English speaking Americans or, whatever, so that's an interesting question i don't know if i'll

get around to that but, it's just a thought for the future.
okay, so

S3: thank you

S4: Speaker information restricted

S1: any questions? yeah

S4: Speaker information restricted

S1: i don't know how much um, i think that's a really interesting point i don't know how much um this question's gonna get to that because a lot of that [S4: Speaker information restricted] was like will you use it for that i mean, [S4: Speaker information restricted] is that a reason for you to keep it up. okay, and in that sense, i guess that's true though because if they say no or they say it's not really important then [S4: Speaker information restricted] for an academic right then that means that they're probably not planning on i dunno studying in it or using it just to do papers

S3: could i just do an accuracy check on this my memory is that Spanish that the responses were much higher for Spanish than Hindi in that (xx)

S1: they're, they're higher than Hindi, (xx) lemme put it up

S4: Speaker information restricted

S1: yeah, here i'll put up this one this is actually, i thought this was harder to read so i changed it to being a, a frequency table but, basically you can see here there's, uh, (wait cuz this is the wrong one)

SU-f: (xx)

S4: Speaker information restricted

S1: okay, this is the one down here... um, for the Spanish sample it's it's still four [S4: Speaker information restricted] point one which means it's important, [SU-f: yeah] but that does mean i mean there's there was definitely several people in the Spanish sample who just circled five five five five five you know, so you know it does mean that more people were circling four or whatever but not, yeah

S5: well, it might be subject to that because um, one of my best friends is actually taking this class. [S1: oh okay <LAUGH>] and she was interviewed. [S1: yes] and i spend a lot of time with her, [S1: uhuh] and all the stuff that you were talking about you know it's this whole, pressure to want to become one who belongs in the community, and to actually, feel that at times because she's not as- she doesn't think she's as proficient [S1: mhm] (then) she might just end up just having this symbolic, baggage (thing.) taking you know (xx) her life that she (xx) become a Sp- Spanish speaker. [S1: mhm] but in terms of the academic stuff, it would seem to me that the class actually encour- encourages the students to look at Spanish in a new domain. [S1: right] encourages very academically based [S1: right] so it would be interesting to actually, perhaps, in another time another [S1: mhm] place, [S3: a new world] <SS: LAUGH> to actually see you know these um language attitudes before the class begins. [S1: right, that would, that would be helpful yeah Jenny mentions] for classes like that and then, see what happens (with it) because um i do know some people who are just totally amazed that, that they can actually feel like they (can) (xx) (their own) Spanish. [S1: mhm] and that they feel validated [S1: okay] in that domain.

S1: okay [SU-f: mhm] so maybe there's some- maybe that has something to do with this difference too because i don't think that that's an emphasis in the [SU-f: (right no it isn't) (xx) emphasis (xx)] Hindi class and when i asked, and when i asked students about that in the interviews um there were m- there were more people i don't_ people didn't mention that they wanted to use it for academics necessarily but but sort of people

mentioned for the Spanish that they wanted to use that for work, um when i asked about Hindi i got kind of different answers you know some people would say oh no you don't even need Hindi in in India you can just speak English, um one person said well it'd be nice maybe if i can do business in India which i'm thinking maybe i'll speak Hindi. but that was like one of the (xx) people i um interviewed so maybe that has something to do with the class i'm not sure, um it would be nice to find that out before sample time but um but it's basically, i think it's true though that Spanish is seen as more useful in this country though. [S5: mhm right. yeah right] at least professionally i mean i think that's definitely even without the class

S3: but there is there is no sign of (a) sharp split on those figures (that were found) in the literature I made this comment before between, domestic functions [S1: yeah] and professional functions. you know (xx) right four five and four one are lower than four eight

S1: right

S5: yeah

S3: and four six but not that much lower.

S4: Speaker information restricted

S1: right i've been trying to find some unmotivated Spanish speakers, but, i have to contact one but but (xx) <AUDIO DISTURBANCE><LAUGH> who aren't organized (then who are organized)

S3: a deadbeat uninterested Spanish speaker?

SU-f: if there are any <SS: LAUGH>

SU-f: we refer to her in one of my classes

S1: yeah <SS: LAUGH>

S1: yeah, i should get some more people from my class cuz i have people in my class who, um, i this_ my original sample is actually my class my (xx) kind of test sample for my questionnaire, um, and i got some, unmotivated people but,

S2: are these all people who are planning on living and working in the United States for the rest of their existence?

S1: all the people? well i asked them that and <LAUGH> what they said was let's see, live and work okay so most of these people are um, U-S born maybe that's helpful to answer that first, um, of all these students i think, five are not U-S born and most of those are, naturalized. so i think there was only one actual like on a visa person in the entire, um group that i got from these two classes. so, most of these students are U-S born, American students, um, and let's see three say that they might that they they will definitely work this is the Spanish sample seven say maybe only one says no for the Hindi, working language A is like three, three and seven more people say definitely no. um and living in a different country only one said yes. someone said maybe, (xx)

S2: did anyone mention travel?

S1: i didn't ask them about traveling i did ask them in the interviews how often do you go? you know um, the Hindi speakers that i spoke to in interviews i mean it's hard to know with that cuz then i i just picked whoever was willing to do an interview with me, um but they said uh, they, seemed to have gone more often actually than the Spanish speakers which surprised me cuz i thought well Puerto Rico is a lot closer and Mexico is a lot closer why not go there? (xx) the Indian <LAUGH> speakers, seemed to have a, closer, relationship one of them, goes like every two years and the other one's been several times and

S2: what about maintenance in order to keep relations with family members who live outside the U-S?

S1: who live outside the U-S? that was what a lot of people said in the interviews they said either their grandparents or whatever their relatives they wanted to be able to speak with them. but i didn't really, i mean i said you know do you speak, language A with your other rel- relatives but i didn't really handle that. (xx)

S3: i think we should begin to wrap up at this point and move on to the next one. otherwise we'll run out of time that's interesting (xx) did you want to,

S4: Speaker information restricted

S1: right um, that was something that um, Holly asked about before here did you ask about it or maybe someone else did but sometimes when i wrote these things i said i don't really know if people think this means you know and that's kinda one of the weird things about doing a questionnaire lke this you know, but it's sort of for the feeling cuz i put no English, uh, i thought maybe that would clarify like that means they do not speak language A at all okay? so that means that it's somebody who's not bilingual, so i sort of hoped that that would clear that up but, (xx) i mean i also don't know what people think about what do i mean by personal you know i mean? <LAUGH> so i mean i guess people can imagine the other thing i put on here for the_ actually for the Spanish questionnaire is is cultural and what do people mean by cultural, (xx)

S6: the first question is this speaks language A only or speaks also English (like) being bilingual

S1: i see i didn't say that, i just thought_ i just imagined people would say does this person speak it at all, and, *what_ to whatever degree.*

S6: *because i think,* could make a difference, *(if you) (xx)*

S1: *if you say they're* bilingual?

S6: if they only speak that one language and you need to communicate with them in that language.

S1: right

S3: right

S1: (xx) so maybe (xx) to ask instead of that

S3: that's the second time that you say that's come up. so, (xx) make it more explicit.

S1: (xx) yeah make it like do you imagine yourself marrying somebody who's bilingual or only language A or... both or or, English or, i could make that two separate questions i guess

S3: we really need to move on [S1: okay] cuz (xx) run into the next i'm sorry Helen

S1: that's okay

S3: i know you were,

S6: my presentation i can, i can give it in a faster or short way depending on the time that is left <LAUGH> i mean it, because it's

S3: well i guess we have to, y- i i guess we'll have to make it twenty minutes.

S6: uh, doesn't matter to me [S3: is that okay?] (xx) i can make it twenty i can make it thirty i can make it fifteen. it's it's it's just the article. [SU-f: (xx) data] what?

SU-f: (xx) data

SU-f: (up or down?) <SU-f: LAUGH>

SU-f: oh i do i do (have data)

S3: well do you want to resolve this by going first? is this_ would that help or what? what's the best way to do this?

S6: you can keep the time that you need and then i fill the rest, because i can change the style

S5: yeah that's fine, it's just that um, i'm not gonna go the whole time so that's fine.

S3: that's fine (xx)

S6: it's fine because i can do it.

S3: so you want an interactive, uh response

<UNINTELLIGIBLE SPEECH>

S3: can we manage to open (door) just a little (xx) it's pretty stuffy or will it just be too noisy

S7: it's okay. can you open it up a little (xx)

S3: just just follow a couple of inches. so i (xx) can get some oxygen. <SU-f: LAUGH>

S6: i'm not gonna use the projector.

S7: okay

S1: sorry it's still (xx)

<PAUSE:27>

S3: do you need this uh thing? [S5: oh no i don't] raised?

S5: i actually need that yes i mean i would like to (just,) write on the blackboard. [S3: right] if possible

S5: it's so funny i have no problems in front of my students...
[S1: just think of us as your students] (xx) in front of my colleagues.

S3: (xx)

<PAUSE:04>

S5: okay

S3: well this matters less less is at stake

S5: i'm sorry?

S3: less is at stake than teaching a class.

S5: oh yes. <SU-f: LAUGH> especially when (xx) an exam. okay i think i need to um sign this really quickly right?

<PAUSE:05>

S1: you can sign 'em after (xx)

S5: (okay i can sign it after) (i guess) (xx) <S5: WRITING ON BOARD NEXT 1:40> hm. okay that's, one thing here (and then there's) (xx) <S5: WRITING ON BOARD NEXT :41 > (the borderline for me,) Native American languages with Spanish (xx)

S3: areas where they're in contact you mean?

S5: yes

S3: mhm

S5: and that would include um i'm (merely.) concentrated on bilingualism and other language contact phenomena. um, how should i start this um <PAUSE:05> let me just give you a little bit as to why i'm interested in this um, first thing um, let's see i um, did a lot of the course work at U-C-L-A on uh American Indian studies. so that's kind of one is my my specialties, um North American Indian studies. and um, i also_ my dissertation interest is in the Andean, um, Andean studies, and um one of the things I wanted to find out was um, what happens in borderland situations in terms of language and culture. and i felt that um, i really had not done a lot of research in this particular area.

i was thinking about doing, my r- my my my um presentation on on Andean bilingualism but i said, i have a lot of time to do (that if i want,) so um what i'm concentrating on is um uh the Pueblo um cultures of the Southwest. and i'm including the Navajos um, separately because actually they're not from from the Southwest they actually came from the north. um and they've been around for like, um... thousands of years some people say um and it's extremely controversial uh in terms of that but they- they're they know themselves (xx) and others know them as not being from there. in terms of the Pueblos, they've been they've been there in terms of their archeological records since um, i would say fifteen hundred years. and um, they've been in (in,) and they and in terms of the contemp- contemporary geogr- geographic locations of these people, you would have to look at, um Arizona New Mexico, Utah, um, i would say what else is the Southwest? [SU-f: Colorado?] i'm sorry?

SU-f: Colorado?

S5: Colorado. yes but mainly uh Arizona and New Mexico. and um... these people have been in contact actually um, with, the Spaniards since the sixteenth century. and that happened because of Hernando DeSoto's um, conquest of Florida um, going you know um, under th- under the um, supervision and tutelage of uh the the Spanish <AUDIO DISTURBANCE> (xx) um after that you have, um when Mexico is reestablished, as a, colony of Spain, you have a lot of conquistadores going up north. so, they- that's when you actually have, the first, contacts of um Spanish speakers, with um, with Native American speakers. let's see... and you have the situation where, these conquistadores actually um get a lot of slaves from these_ from the- from these communities, and they actually, teach these people Spanish so they can serve as interpreters. so you have a whole line of families, involved in this, and from this um, y- um you have the spread of um, of, translators and also you have the spread of missionaries because, along the with the conquistadores came the missionaries. so, the medium, as is in the case of the Americas, the medium for, um proselytization is, um, Spanish. in terms of

um, the s- the um the Jesuits and other um Spanish uh missionary people... you have um, you have the English speakers coming in in this century which is really interesting you have them coming in for mainly commerce with the Pueblos and also for expropriation of their lands so you have a whole, a whole historical struggle, of um, of um Pueblo, peoples and Navajos being transplanted, from dif- different parts of the Southwest kind of being transplanted from one neighborhood to the other and this happens, and by neighborhood i don't mean like you know small, localities but actually wide regions. um this happens actually, um, throughout this century um even now in the Seventies you still have conflicts over land. between um the Spanish speakers um the mainstream, um political authorities, Congress, and um Pueblos and Navajos. um, you also have a situation, where, because of all this, linguistic and cultural contact, um you have a lot of loss in terms of both peoples and languages. and the three major languages of this area are um, Navajo, Pima-Papago and Apache. which i was surprised because i wasn't_ i wa- i i didn't think that Apache was uh, actually, the d- the dominant language because i felt that_ i thought from the history that i know that Apache speakers were actually decimated. in um, in the late eighteen hundreds. where you still have um, mu- monolinguals and bilinguals especially. um for instance you have the U-S-A Bureau of Indian Affairs citing um um, a- an approximate number of speakers for each of these um languages you have about, a hundred thirty thousand Navajo speakers, you have, fifteen hundred, Pima-Papago speakers and you have nine thousand, Apache speakers. they really don't delineate exactly, who these speakers are you know in terms of the bilingual continuum. but um i guess y- you can imagine that a lot of them are bilinguals. um within the Navajo nation, which is a nation which is a which is amazing nation i think, it actually covers four states that's why it's called the four corners. right? s- and and that (happened) as a result of all this trans- transplantation. as a resort_ as as a as a result of them moving, as soon as you have, um, Anglo speakers and also other Native American communities you have Navajos being

transplanted, both forcibly and also in terms of you know in terms of their own desire (xx) wanting that, those um, (dislocations.) um... and and when it and when it come to dealing with Native American languages and Native American speakers in the Southwest, one also has to deal with Chicanos because both of them, Native Americans Chicanos are actually what you consider territorial minorities. they're actually people who if you have to say who were there first. and within those two communities there's a lot of controversy as to who was there first. but we don't want to get into that. we just need to know that um, let me see <PAUSE:06> you have another situation going on is interesting. you have the situation of the Southwest- e- uh Southwest and Mexico. um, several researchers uh researchers including Jane Hill, have uh conducted studies in loan words. and they've found e- that um, you have, you have a considerable degree of um borrowing, um in some Native American languages, um, they've found Nahuatl borrowings. which come from, as early as, you know, i would say even before the conquest. and you still have those remnants which is really interesting, um... what's interesting about the Southwest too in terms of the populations there is that, you don- you don't only have bilingualism you also have trilingualism, which becomes much more, complex in terms of studying, um, those linguistic um, repertoires. um, you have situations where, um usually in terms, there's certain typologies of language contact in native North America but, n- in focusing s- the southwest, you have speakers, who, speakers who are um beyond thirty, usually are bilingual. in both, n- n- bilingual trilingual, um in their Native American language, um Spanish and English and you have speakers who are younger of course usually are much more English dominant. and in this situation you have a situation of um, um this phenomenon called, Indian English. which is uh which is very it's been categorized very similarly to Black English, or Puerto Rican English where you have um certain um phonological um intonational, s- semantic um, similarities from the, from the native language. um, you have this this scholar by the name of William (Leap.) who's actually spent_ he's actually um devoted

his his life to um this type of um, of study, and he usually has done it um, he usually has looked at um... Navajos Navajos

S3: this is specifically uh Navajo English [S5: no] he's interested in or is it the whole uh repertoire

S5: the whole repertoire what's interesting about it he he he, that's a really good point, he does he does adm- um, he does state that, as there are different types of Englishes you also have different types of Indian English so you have Navajo Indian English you have Hopi Indian English you have Cherokee Indian English. so, one has to also get to know the native language in order to figure what are the similarities and whether they diverge. so i i would say the same thing with Spanish. you have Puerto Rican English and you probably have Colombian English. although, uh depending on the on the contact situation you could have, you could have dialect leveling in that sense too. um, the history about this Indian English is interesting because, by the turn of the century and this was like a, widespread um, um policy for Native Americans, by the turn of the century you have, you had in- um Indian children being forced out of their homes into boarding schools. so you actually have, um kids coming from different s- uh reservations, different places around native America, who actually had a had to deal with each other, and um, and understand each other so English became like a lingua franca. so within that you have that situation going on you also have, in the nineteen fifties the termination policy which actually, um, expelled a lot of the native peoples into the cities and that was another, governmental policy, where you have people also coming from different, um linguistic backgrounds different languages, um, coming together and using English as a lang- lingua franca. but then you have the situation where people become, dominant and people actually, just, using this Indian English. so um there's been some research there's research on that but still it's ongoing [S3: mhm] because it's still you still have, so many different languages you still, the typology's still not there but uh it's it's a a field that is very very um popular right now.

<PAUSE:06> and what's interesting about it, i- in the situation is that (this i- this i didn't) find in the literature, if one has Indian English you want s- (you want suppose they have Indian Spanish) but that is not, that really has not been seen in the literature, i mean i think it's to do, you have a question yes

S2: do you get cases, where there're Native Americans who speak Spanish but not English?

S5: yes you do you actually have the older f- uh the older people, in certain par- in certain parts of the Southwest. there're some older uh generation Navajos that that are actually, bilingual only in Navajo and Spanish. (xx) depending on the closeness of the interaction with other Spanish speakers especially Mexican, speakers, you have the much more Spanish dominant. but within that you know you can still talk about Indian Spanish, just the same way you can talk about, you know, the Spanish that Indian Spanish that's influenced by Quechua or you know, other languages Native American languages. so that's that's interesting i guess it's because we are in an English dominant, society academic world so, i'm sure other p- i'm sure people are studying that too. um... let's see what else i can, tell you guys. there's there's been one, interesting study and it's it's finally, we probably will be very appreciative of it. has to do with um, with uh Susan Phillip's study in nineteen seventy-two, of this um war- Warm Springs um Indian um classroom, where she actually um, studies the interactio- the interactional dynamics of um Indian children. and how they use they use silence, how they use, how they use, uh the notion of uh of group solidarity, to actually both both um, create a solidarity and also separate themselves from others. so you have you that and i believe that she, um used a little bit of the (xx) approach but not, the way that we've been doing it much more (xx) and this is kind of like a a a very um, good field to get into because um, one of the one of the biggest biggest urgencies in terms of this research is educational educational um um, stuff. um we have in terms of Indian English you have,

the same old thing about the literature by semil- semiling-
lingual speakers, the same things is appro- is is attributed to
um Indian English speakers. um, and you also have the whole
situation not just linguistic and this is interesting too it's
like not just, y- you, the the um, the person, specially the
academic child is not just a a linguistic person in terms of um
speaking. (actor.) but also you have the, the situation of
gestures. how gestures are read. are read by the people, in
authority and how gestures can vary, cross-culturally.

S3: well this is real ethnography of speaking [S5: yes] stuff
that we've already looked at i think in this class

S5: exactly and [S3: s-] that that is something you know [S3:
it's (the Dell) Hymes line] mhm, s- and with a little bit of C-
A stuff there. so, and that's actually the only research that
actually has been done, i'm, i'm sure people have done more of
this type of research in terms of uh um educational typology
perhaps, but in terms of actually being a research that goes
beyond the educational classroom, that actually piques the
interest of other scholars this is, this is a classic one. and
that's and that's going to, um... you have let me read you have
some some sit- some um, some language attitudes from different
people in... among the Pueblos. let's see if i have them here.

S3: you'll need to wind up very soon, (xx) [S5: McKay this will
be my last thing] (xx) we'd a slight screw-up (xx)

S5: this is something that um... is very very interesting. and i
hope i can find it. <PAUSE:04> maybe it's here <PAUSE:16> that's
too bad i didn't bring that (xx) um if you guys have any
questions, feel fr- feel free to ask, (as i look around...) mhm?

S4: Speaker information restricted

S5: in some places it is i would say in some urban se- sec-
sectors, this happened because some, some urban, community
places have a longer history than others.

<AUDIO DISTURBANCE>

S4: Speaker information restricted

S5: besides English is (xx) and Spanish in some places but usually s- Spanish comes secondary to English yeah English is totally taking over the linguistic repertoires of (Black) communities

S4: Speaker information restricted

S5: well i don't know i mean tha- that's, <LAUGH> empirical question again

S4: Speaker information restricted

S5: yeah yes it's so little research actually, and i think it's i- it's just because of the complexity of the linguistic and cultural situation, um, one really_ okay this is one thing i would like to end with which i thought was (xx) really interesting there's this article by (Michael Silverstein.) um, and let me, tell you a little bit of this article, if i can find it here. okay this article is in um Journal of Linguistic Anthropology and he has an article Encountering Language and Language (of Encounter,) in North American ethnohistory. and one of the things he does says is that if you're gonna look at a language contact, phenomenon in native North America or the Americas when dealing with n- with uh Indian languages, um one really has to look at the histories because, there's so much history of transplantation of removal. so the people who you at this point think are, the, native speakers of that region might not be at all. and this could just be, fifty years, you know, down the line, doesn't have to be like hundreds of years down the line. and that's one of the things i think when you know that's one thing at times that we might forget to do to actually do our research and not that our research will will necessarily um um, give us a one-to-one correspondence in terms of, how language, is being either bor- borrowed or not borrowed or uh

code-switched, but it will definitely make our_ the complexity much more realistic.

S3: thank you it's very very interesting

SU-5: thank you for (xx)

S3: thanks very much (xx) *<AUDIO DISTURBANCE>* okay so just just take your time.

S6: okay, what i'm... what i'm doing today presenting today, it's n- nothing, more or less than the article that we had from this book that we talked about in class we started to talk about in class, and, we_ i promised to finish. and the title is Code-Switching in Bilingual First Language Acquisition. and i'm going to be following your handout so you_ if you get lost or something just stop me.

S3: would you like people to ask questions as you go through it?

S6: please do

S3: okay mhm, this is an important article so take advantage of her superior, knowledge (xx) in your syntax class

S6: *<LAUGH>* well, i tried to, to concentrate maybe more in the parts that we'd been doing in class and less in the syntax, [**S3:** mhm] but let's see if we, we_ we can, make it together. okay, first thing to start with is talking about the, terminology, because when we talk about bilingual first language acquisition, uh Koeppe and Meisel, these authors that are by the way German working at the University of Hamburg, they refer to children as opposed to, an acquiring two, or more languages simultaneously before age three and this is very important because, we've seen the differences before. so (there's) simultaneously, both languages at the same time, before the age three. and these children could have lived in one or the other country, by th- until the a- age three comes. um, what they talk about is bilingual code-switching, which is governed by grammatical and

pragmatic constraints which means that code-switchers, are required to have pragmatic and grammatic, and grammatic competence, grammatical competence in both languages. um, but what happens with infants is that sometimes bilingual children use both languages, in the same utterance or a conversation, and they violate syntactic or grammatic constraints of code-switching, this is what they call code-mixing. and, it can be accounted, for by the failure in separating the two linguistic systems, that's what is called fusion, due either to the lack of lang- of knowledge of syntactic and pras- pragmatic constraints, or also could be to the lack of, the necessary elements to which the constraints applies, or maybe both. so there's, one thing they're gonna have a look into. um, important point is to determine from what age children organize language by grammatical means. and also, um the distinction between different categories of words and having these two elements, um, counting for their, study, helps them to distinguish code-mixing code-mixing from code-switching, in early child language. so it depends really, on, from what age onward children organize language and when are they really, using their syntactic knowledge or not so that's that's really important for this distinction between code-switching and code-mixing. and what Koeppe and Meisel do is um, comprehensive literature review about code-switching acquisition, and they, they say that these studies are mainly referred to longitudinal studies on bilingual language acquisition which typically, concern a linguistic environment, where each person speaks one language. so in all these situations so one parent one language, or child_ children speaking one language in the home and then one language outside of the home, they said that most studies, they referred to this. which is true i've seen that here in class. um, they divide their, their studies in two, um main (tracks) like we've seen before one is the pragmatic functions of code-switching and the other one is the syntactic aspects of language mixing. and, starting with the pragmatic functions we've seen two, two types in class do you remember which one they were? do you_ can you, just more or less a rough idea wha- which two aspects big

aspects talking about the pragmatic functions of code-switching that we've seen in class?

S1: to change the situation, somehow, to indicate change of situation or in, context [S6: mhm so] like formal informal

S6: exactly one was situational switching, and remember the other one? one was a situation and the wu- other was one, was the

S4: Speaker information restricted

S6: the topic situation and the other one was the speaker. mhm you know Gumperz and everything we've (been through)

S4: Speaker information restricted

S6: okay, well, really this_ about situational switching, language choice is, the ability to select the appropriate language, as base language for the conversation, according to the interlocutor, which is uh (oh) the topic, of the conversation, or the situational context. and from around age of two, the child switches ad- adequately languages according to the interlocutor. most children raised bilingual in a monolingual environment could experience <AUDIO DISTURBANCE> that's why, it calls their attention of Koeppe and Meisel. that means this language factor that means that the fact that most people talk in these children, um, most people who talk to them only understand one of the languages, um, these children start to develop a a certain awareness like, i'm talking, to this person and this person doesn't understand me, so it means, they talk something else but i understand them maybe i talk two things. i mean this is like, the way ch- children could think. and this, to that they become more aware that there's a language separation that there's two languages going on. this factor interlocutor becomes complex especially with bilingual interlocutors because children need to decide, which language, they're going to speak with this person and, which language do they prefer or they_ if they prefer to use both codes at the

same time. um, the other function of code-switching we talked about is, um... the_ apart from situational switching we have also conversational switching. which is the other, main um type of, of switching, and in case of bilingual hearers, only, code-switching may be used to convey specific social or pragmatic information, such as metalinguistic awareness. metalinguistic_ sorry excuse me metalinguistic comments that reflect their awareness. this is called like i said conversational switching, in contrast to, language choice as situational switching. do you remember all those cases in which, in the middle of a conversation, people will switch from one language to the other we've seen it, (a lot of them?)

S3: we also had it modeled according to Auer's definition [S6: mhm] if you recall between participant uh oriented switching, [S6: mhm] and discourse [S4: Speaker information restricted] [S6: oriented] oriented you know these, [S6: mhm] and that's the most recent version of that [S6: mhm] binary category you're talking about

S6: mhm, which is is basically the same, with a different terminology, yeah this this binary, category. well what is being reported is that from age two onwards children begin to repeat their own utterances, in both languages in order to call for instance the parents' attention, or to ensure that they're being understood. from the age three onwards they start to comment on their own language and to ask for translations, and this is really a sign of separation of languages. and from age four onwards some children are reported to use marked language choice, as a means of amusing the hearer or include a third person from the conversation [S3: mhm] we've seen this as well, before [S3: mhm] and we'll see some examples for that. um, with respect to the syntactic aspects of language mixing, uh Koeppe and Meisel they gathered the, the literature, about, the studies that they, they're concerned with language mixing and language separation, the problem is, most of the studies find out that the mixings are very low, we've seen this in class as well. this is m- mainly repetition of many things

that we've seen so far it's a bit of, um, sum- summarizing. what is necessary is to relate mixing rates to polititive aspects, of language for instance which categories are mixed, the structure of condition for switching how developed is the child, is the child's grammatical competence etcetera. and also to pragmatic functional considerations, are these mixes conscious? do they respect social rules of language use? is the addressee bilingual...? the distinction between different categories of word mix, helps to distinguish code-mixing from code-switching in early child language so we need to distinguish which categories have been mixed, to start with, and empirical studies have shown that in early stages children mix function words, what are function words? do you all know? can somebody explain that?

S4: Speaker information restricted

S6: mhm mhm well yeah, all categories except nouns, adjectives and verbs. which means they're, they're more like function words, tha- words the words that have a function but not really a content (xx) they're not in the dictionary.

S5: (xx) differ from content [S6: mhm] versus function

S6: content versus function or lexical words versus functional words f- in the fourth page you have here, uh the distinction, in the f- last column, of the mixing, tables, there's single words mixed German and French and they distinguish between function words and lexical words. and the percentages that we're gonna talk about later. okay, um, was an author called Wieman who talk about, these function words being mixed in early um, stages of of children bilingualism, the problem is this is very, um, problematic, for the terminology and also, because there are many contradictory studies about this, so people keep um, finding, contradictions and, Koeppe and Meisel said this is what has been found so far but they found, they found, pretty much the same tendency, but some studies like for instance Lanza's study Elizabeth Lanza who was, um, published the last review in

ninety-seven, she found contradictory findings about this it's not exactly, um, that children start with function words but that's a tendency, so far, it's been observed, and later on, nouns are predominant, it's been found that seventy-five percent, of the switch words are nouns in both, adults and children. now this coincides as well with borrowings most borrowings are also nouns.

S3: did we ever discuss why nouns seem so, you know there are id- people have ideas about why nouns, uh are so liable to be shifted about okay? Mary you're nodding do you know why?

S7: um, one of the things i read was that there were more nouns, [S3: even] like [S3: even just] in the corpus for example i think Poplack [S3: yeah] mentioned, nouns are switched more (xx) oh okay

S3: even allowing for that though i think there's more more nouns are transferred [S7: oh] than in the language as a whole [S7: yeah it's not] there're there are syntactic re- well. i mean basically the argument is they don't drag a lot of syntax around with them <SS: LAUGH> you know if you have a verb for example some verbs are transitive some intransitive, [S4: Speaker information restricted] [S1: yeah] if you take a verb like um, eat, eat, need not take an object but if you take a verb like devour, you can't say John devoured you've gotta say John devoured something, and so if you transfer that item to another language you have to know, uh this a- you have to have this abstract knowledge, about how to use it, and that's uh that's it's, well, i call it subcategorization you might have another way of putting it

S6: subcategorization <LAUGH>[S3: okay] exactly

S3: but you see nouns don't do this, nouns are relatively easy to move about they just don't drag the syntactic complexity, it's to do with the verb being the head of the sentence effectively really isn't it? [S6: mhm] just having this central role, in the organization of the clause?

S1: (well)

S5: and that would be cross languages?

S3: apparently yes. uh s- i ha- Twaila Tardiff was giving a paper in Chinese and suggesting it didn't happen with Chinese, but i'd've liked to look at her [S5: yeah] methodology because Chinese um, has strange rules for anaphora [S5: mhm] you know you can actually cut all kinds of things out of, Chinese utterances and connected speech but it's it's a pretty solid finding, it's coming up again and again.

S5: maybe something about Chinese also

S3: yeah, well, it's just that we we know very well, that people don't like violating syntactic rules of languages whether you mix and it's quite easy, not to do this with ba- with nouns but if you m- move verbs about and some other things like prepositions [S5: right] it's much harder

S6: mhm, it's true. um, that's, i'm gonna retake up the idea in a minute, um generative grammar tried to explain formal regularities in code-switching by means of universal principles. that's really Chomskyan studies, try to find universal, principles under, linguistics under every language, although there are contradicti- contradictory findings it is agreed that mixing seems not to occur between constituents that are contained, in the INFL phrase. um, i don't know if you know what the INFL phrase is. inflection? um it's for instance um well, in a verb, we have inflections m- the ending of the verbs, in Spanish for instance the verb hablo <WRITING ON BOARD> i speak, we have, um, this O, tells us that it's first person singular, um, and tells us as well, that's it's a present tense, that's inflection, in the verb. so y- wou- wouldn't find things of the type <WRITING ON BOARD> i don't hablo you wouldn't find these kind of mixings because, in the INFL, f- phrase, and also, like things like no quiero <WRITING ON BOARD> to talk, we wouldn't find this at all as well, um because it's between um for instance nega- uh, yeah, a finis a non-finite verb, you wouldn't

find this [S3: right] mixture, you wouldn't find it between a subject clitic, and a fin- and, and a finite verb, i wouldn't find_ i don't hab- oh sorry i hablo <WRITING ON BOARD> and also between, when, a negation and a finite verb, you wouldn't, either find this switch

S4: Speaker information restricted

S6: i hablo?

S4: Speaker information restricted

S3: this is the whole problem with this, [SU-6: exactly] i was just waiting for (xx) <SS: LAUGH>

S5: the first the first one's kind of like, the first one's kind of like mock Spanish, to me

S6: yeah i don't [S1: (xx)] hablo [S5: yeah] espanol <LAUGH> something like that

S1: alright i don't know

S6: it's true , it sounds like mock Spanish, well what they found is in the data they_ these shouldn't appear these, words, m- judge as speakers as non- not, nongrammatical.

SS: mhm

S6: (xx)

S3: judgments are one thing what they actually do are another as you know <SS: LAUGH>

S6: exactly. i recently gave a presentation like i said about this topic, and it was also (auto) code-switching, and all the authors, they were based on data, and, judgment, and they got always contradictory, findings, from one or the other, and they, they weren't sure they just guessed, distribute this way and, there're too many contradictions

S3: i think there's no pro- you see we did discuss this whole sort of whether these constraints were variable or categorical [SS: mhm] if you accept there's a tendency, [SS: mm yeah right] to (avoid) that i think you're not in trouble, [S6: mhm] but if you try to say it's categorical [SU-6: (xx) yeah] then you probably are

S5: i think it's true when you mentioned that in that meeting with um with um, the woman who does the language contact stuff?

S3: ah Sally Thomason

S5: yes

S3: yes

S5: it was very good that you said that because i think, she also accepted it i mean most people would accept that because that would be common sense to actually say that, instead of saying that it's categorical because then you have (xx)

S3: well as you know theoretical linguistics tends to work categorically [S5: yeah] it doesn't like operating [S5: right] varia- you know with these variable constraints

S5: (it doesn't like) sociolinguistic tendencies <SS: LAUGH>

S6: mm well, w- what we have said about the INFL phrase so inflection is, what i- it's what they call functional head, and specially the f- well this functional category inflection, is what it seems to constrain code-switching.

S3: right

S6: and l- we said, uh noun phrases nouns they don't present any any constraints they are freely moved. um, research concerned with acquisition of code-switching, and syntactic constraint has to consider, for children the gene- general development of syntax in both languages that's really [S3: right] important... um, <REFERENCING HANDOUT> the study that Koeppe and Meisel

carried out is called a DUFDE study i guess DUFDE because it's German <LAUGH> um le- it's a longitudinal study of simultaneous acquisition of German and French by, thirteen pre-school children. they ve- videotaped the children, every two weeks for (thirty) minutes in each language, and then they transcribed both linguistic and non-linguistic interaction, with their context. and i think that's very interesting because most people do just, audiotaping, and they just get the speech but they also did the interactions, the non-linguistic interactions, in the context to, really to catch, the situation especially for, um, conversational switching was important. the results presented concern pragmatic and syntactic aspects of the of the (mixed) speech of two children, Annika and Ivar, both first born growing up in middle class families in Hamburg, both mothers are French and both fathers German, and each parents uses her or his respective language to talk to the child. Annika's parents, talk to each other in German where, Ivar's parents talk in French. French is initially the dominant input language for both children because of their mothers, but after age two, two w- two six two six means two years and six months, um Annika's German was stronger, whereas for Ivar both languages were always in balance, and that was due t- can you imagine, what was due to? what was depending on that one had German very strong and the other one both in balance?

S1: (was it) in the homes, which one was spoken in the home i can't remember now

S6: mhm Annika [**S1: which one?**] Annika was speaking German in the home

S1: German in the home [**S6: mhm mhm**] okay

S6: yeah so because of the social interactions travelling to, travelling to, to France and coming back and visiting friends, uh one developed higher, stronger German and the other one was more balanced. and during the recordings the interviewers were

supposed to keep the languages separated, it's all this issue with, yes?

S2: do they mention anything about the development of their accent, mm at all? **[S3: mm]** i guess they're young enough so **[S6: yeah]** it wouldn't make any difference never mind

S6: mhm no they didn't say **anything**

S3: **actually** there was question on the list on the on this dreadful linguist- uh bilingual list on <SS: LAUGH> email that we're all getting it's seventy messages a day (xx) somebody was pointing out there was a paucity of studies of phonology, of bilingual children i think hardly **[S5: yes]** **anything** is known about it

S6: mhm it's true

S5: i saw that, and it did say that.

S3: uh, i can't think of a single one actually

S6: what is, is amazing is children have such capacity of reproducing sounds even though they don't know what they're saying, i think

S5: (xx)

S1: the only thing i was gonna say is all those students that i talked to in interviews i asked them like, they've often said like um, they_ like the Hindi speakers they said like my Hindi is horrible but i can say all the words exactly right **[S3: right]** (xx) have a good accent. and the same thing as the Spanish speakers they said well i have a really good accent, cuz they would always take Spanish in high school, and they they'd all talk about how they weren't that good but, they could get by cuz of the accent and stuff like that so it seemed like even though their Spanish their language wasn't that good the phonology was there

S5: well it seems yeah

S3: sorry, go on

S5: oh okay i was gonna say that um, i asked when i was taking phonetics i asked Pam Beddor, [S3: mhm] about, if she knew of anyb- uh research in bilingualism and phonology and she said that there's really it's, yes

S3: no i think there's, there's very little, but that's a good point you raise because it's very salient to speakers um i (xx) a very similar experience i used to speak German reasonably fluently but it, went away because of lack of practice over about twenty years, then i went to Austria, and i was in um the Tyrol where they speak a very very stigmatized dialect very non-standard, and i, staggered a few sentences together in German, and the chambermaid said ah klassisches Deutsch <SS: LAUGH> wonderful German, and that was my accent i'd been taught to speak standard German. and, they were ashamed of their, accent and she was prepared to ignore, the awful things i was doing to the syntax. <SS: LAUGH>

S3: just like, you know a bit similar i think to your Hindi speakers, but it's a it's a very very interesting, area i don't think anyone's looked at it. so, more research ideas.

S4: Speaker information restricted

S6: oh that's that's really the, the um, la- the last table in the last [S4: Speaker information restricted] page [S4: Speaker information restricted] because they they mention their utterances, [S4: Speaker information restricted] um, they don't really explain it that much, how they measure, but uh because it's just an article i think you should need to, you need a more, a better [S4: Speaker information restricted] (xx) thorough help. and i think they made sure in the total of utterances when they recorded, [S3: right] one of the other, they measured, and also error i (xx) suppose

S3: yeah predominance they do explain i think what they mean it's okay to use terms like balanced if you say you mean by it <LAUGH>[S4: Speaker information restricted] and i think they do say basically that they use, the languages more equally, just in terms of quantity, [S4: Speaker information restricted] i think it's that more than error in fact

S6: mhm

S4: Speaker information restricted

S6: okay... we continue, um <PAUSE:04> it is important to separate both languages, for one reason because um, the communication will be determined by the automonolingual, that means that, code-switching is expected to occur less often in a situation that is pure monolingual, than in a bilingual situation, but it's interesting because the use of two codes will be motivated by situational, uh, factors like language choice according to the interlocutor. so if you have a monolingual pure monolingual context, and mi- mixes occur as well, it will be interesting to see why th- why do these mixes occur there, uh as o- opposed to if you have a bilingual situation when you have speakers of both languages, it would be more logical that children mix because they're talking they're code-switching they're talking to both, um speakers but if just now pure monolingual, and they mix this is, this is really, um a chance to observe the motivations, and the external factors that children may have to mix these languages in this situation that should be monolingual. the empirical results of the study, um, they said that's at an early_ starting with the pragmatic functions at an early stage of development if the child's utterances do not correspond to the regularities, um in use, in its linguistic requirements, we may conclude that the child has not acquired those rules, but this may not be true, because it could also be that the child has developed his own rules, and, of language choice and then they don't cor- coincide with the ones of his environment. so this is the two possibilities they're playing with. in, in the data both children they select

the appropriate language with each interlocutor from age one four, for, um Annika and one five for, for Ivar onwards. and the first code-switching appear when the children were addressed by another person and answered to the respective language, we have here in, on the third page you have in the handout, you have an example, (i'm gonna try) with my French, *et comment est-ce qu'il va partir a Paris* um says the French, um, interviewer *avec le* and then, and Ivar directs himself to the German interviewer and said *will weggehen*, i mean so he's switching, uh languages according to the interlocutor. um, su- switches occur fr- like we can see here from, two ele- he is at two eleven but it start to occur from two five, onwards and even two years old, sharp for Annika. [S3: mhm] um,

S3: it's amazing actually

S6: mhm it's amaz-

S3: a lot of people claim that bilingual children have much better developed metalinguistic awareness than monolinguals [S6: mhm] (this) kind of [SU-f: mhm] rings a bell

S6: also self-initiated switches are really interesting is when the child addresses someone without being asked, and started to appear from age two point, two eight, for Ivar and two, for Annika Annika seems to be pretty, uh advanced for some things but we'll see later, that, there's some contradictions in her as well, and there's an example here, example number two when I-Ivar interacts with both interviewers switching between languages and translating wi- and also few errors occurred. uh s- the German interviewer says *<READING FROM HANDOUT; ALTERNATING SPEAKERS> frag sie doch mal "was isst du denn gern?" eh brot du pain, de pain hein quoi? nee non de pain comme manger. he corrects himself he says no no no, no, uh bread no no e- what i mean is like, what you ea- to eat, so he corrects himself, so he's aware, of the switch in languages of the the person he's talking to... and they're self-initiated switches he, um, it came from from himself. there's an example number*

three with Annika when she corrects her errors and adopts different strategies to avoid errors, or when she uses the incorrect language. um, the German, mm interviewer says <READING FROM HANDOUT> Annika hat uns gerade erzählt das sie selber der osterhase ist, and, the m- the mother says c'est quoi? and Annika says to parce que eh and then, she, hesitates and, turns to the German interviewer and says wir machen die eier. so she, um, she realized and and corrected herself. um, from age four, four onwards, Ivar uses marked language choice this is something more elaborated, e- in order to exclude a person from the conversation, and this situation is um, she was, talking to his mother in German, and switches back, to French, u- to sw- to let the French assistant know he doesn't care that she underst- if she understands what he's saying, in German. which is really curious because, we know that Ivar talks in, French [S3: ahah] to his mother, so he run- he really wants to enx- exclude this person as he's using German consciously. <READING FROM HANDOUT> boese kleine maennchen und dann sie m- ham sie aber bestimmt angst. non je ne comprends pas, oui mais ca c'est de l'allemand dis donc, elle comprend pas Marie Claude tu sais oh ca fais rien <LAUGH> it's it's just really rude saying that i don't, i don't care really, <LAUGH> so he's being rude on purpose but, he is als- al- he is already four years four and four months (xx) he's older, now he he, he went over those three years of first, hesitations as some of us call them, and now he knows exactly what he's saying and his choice is exact for the situation. um, talking about the syntactic constraints in code-switching, um it was observed a high mixing rate, from the beginning of the records, onwards, which at the ages of, two five and two zero for for Annika, in- decreased. uh with respect to the frequency of specific syntactic categories like we said nouns and noun phrases are mixed throughout the whole period of the investigation but function words decreased, constantly from age two fou- two five and two, respectively, which this coincide with st- previous studies, and concerning, the sentential position of switches, um after two four most mixes concerned sentences with a mix of elements and can occur between all kinds

of constituents, as we see in example five, *das bateau*, is_ <LAUGH> we have it between the, uh the determiner and the noun, *nounours* and we have two nouns and *tombe berg* we have a verb and a noun... and, this preference cannot be accounted for in terms of syntactic constraints, um, because, well we h- they assume that they don't have syntactic constraints at this age yet. because they_ the category, INFL hasn't appeared yet.

S3: right

S6: from age two five onwards, violations of Poplack's free morpheme constraint, were of_ not observing Ivar's speech. um, i don't know if, you remember, um, Poplack's free morpheme constraint i can remind you if you want to, it says that no switching can occur between a bound morpheme and a lexical form, unless that phonological form is, phonologically integrated in tha- into the language, mm to the bound morpheme, for instance you remember, we have an example it was *catche- catcheando* <WRITING ON BOARD> (xx) from like catching and, the *ando* is the, the continuous gerund form in Spanish [S4: Speaker information restricted] you you have like <WRITING ON BOARD> *flippeando*... like flipping *flippeando* that's okay because it's integrated, but you can't have *catcheando* [SU-f: mkay,] although, we've s- <SS: LAUGH> heard them before right? yeah that was my argument too my other presentation i said, this is what Poplack, Poplack says, this is what i have heard, i'm wha- this is what i've heard, from speakers. so it's, again... one thing is reality and the other thing is what, depending on the data you get, um, an example example six reflects these violations of constraints in Annika actually. um... because um, well she has a few contradictions with respect to, to Ivar. and, is, it says *nounours* oh, excuse me, no no, yes f- no this this is, this is in Ivar's but it's it's earlier and later on he does it from two five he doesn't do it anymore, this is what he used to do it *nounours* *il a reite* and *deddy resucht* so it's it's a mixture between, French and and German, like *reite* is from *reiten* in German, but with the ending of French. [S3: right] mhm, and *reis- sucht* *sucht* is German for, for seek for look for, and this

re is the French, uh, (part,) um... after the age of two five again with Ivar, um, Ivar's (xx) (missing) was reduced to insertion of single nouns or switching between determiners and nouns, in the object noun phrase. we have that, we can see that in example seven, moi je va a la kuche, so it's uh, just the word kuche kitchen inserted there, uh, and like i said before Annika's data is con- contradicting Ivar's data she uses the switch within sentences with omitted elements at a later er-age. when she was already four years old like we can see in example eight she's saying il a gewonnen and also ca c'est Daniels we talked about this in class she's, given the word_ what do we call uh, gen- what the genitive in German she's attaching that to a French in a French sentence. and i've s- observed that in Spanish speakers s- Spanish English bilinguals they do that, i've -served that at age of two two. and, at the age of three seven it's supposed to be, eradicated but, Annika said it. <SS: LAUGH> so... Koeppe and Meisel attributed that probably to performant errors because this type of data was infrequent they said. <SS: LAUGH> the problems they said well if our theory, is consistent, then it's impossible that at this age she has not acquired syntactic regularities that means it must be performant errors and then we wash our hands and save ourselves

<SS: LAUGH>

S3: well a lot of people think that code-mixing is a performance phenomenon, [S6: mhm] you know i- i don't know if you've discussed this with San, [S6: um] or with [S6: yeah] anybody, but, uh this is the huge problem and this is why, if you use variable constraints in fact it at least accounts for the data but it still, doesn't tell you what's happening underlyingly.

S6: mhm. yeah that's the most difficult thing [S3: mhm] to see, um, well continuing, almost finished, um on the basis of qualitative and quantitative changes it seems to be two stages of syntactical development in children according to all this data and, um, and one will be characterized by the absence of

the functional category infi- INFL the infla- inflection, which doesn't appear to be syntactically, constrained the the utterances that that means this co- code-mixing we talked about, and around the age of two and half, children stop mixing function words and rarely violate syntactic constraints on code-switching of the type Poplack, uh presented and this coincides with the emergence of the functional category INFL. which seems to account for qualitative ch- changes in the child's speech by the transition from mixing function words to lexical categories nouns and the ability to respect grammatical constraints in code-switching. in Annika's data, from age two six onwards most of her mix- mixes seems to concern single nouns inserted and the switch, um, at the switch point dertermina- determiner noun like the ones we've seen, before. and if we see this um, um, tables, we can see the tendency the the utterances in German what he says, um, total mix_ we can say we can go for the percentages maybe it's the easiest percentages of mix for German utterances and for French utterances, we can see that, there, tend to disappear towards age three, and we compare, mm Ivar and Annika Annika is, two eleven she's practically also three. and if we take the very last two columns in both um, tables, we can see the function and lexical percentages [S3: right] for the switches, and they're, they're going, the functional they're disappearing, whereas the lexical are present. there so hundred percent, from age, three five, approximately for one child and, yeah two six for the other, there's a clear separation. and we can see that all switches are lexical and function switch-functional, switches functional word switches are are gone. which this uh, corroborates, with the first assumption and other, other authors' um hypothesis too. uh the conclusion, is that changes in formal properties of the mixed speech of bilingual children are related to grammatical developments. [SU-f: mhm] the the completion. the appearance of the functional category INFL again accounts for changes in code-switching. the development of pragmatic function however the the the doesn't seem to observe grammatical constraints, they're separate things after all, although they're inter-related, but we're talking

about different things, pragmatic (aspects of) code-switching are influenced by several factors of language use in the family the parents expectations, about language choice and also the reactions to code-switching. um, which may contribute to the separation of languages and the development of specific strategies to help (avoiding) mixing like with Annika, (xx) points for further investigation, uh will be, to have a look at the causes of early mixing, are they pragmatic motivations for mixing of function words or does this kind of mixing indicate fusion of both language systems, this eternal conflict between, are they two separate systems from the beginning or only one? and i think this is still under debate. and also correlations between grammatical developments and formal aspects of language mixing, are they really universal? and that's why, Koeppe and Meisel suggest that, it would be interesting to analyze the the speech of more children acquiring different languages with respect to these questions. and that's where, this study i've talked about before Elizabeth Lanza, she she does thi- this kind of study with Norwegian and English, and she said that it doesn't hold out this, k- this Koeppe and Meisel what, they, they found out it doesn't hold up, their in her, uh informants, there're changes there're different, um, well results, with respect to this analysis. and, she found out that that her informants mixed function words until the age of two seven. so it's a lot later than than (xx)

S3: well they did actually stop doing it?

S6: pardon?

S3: they did eventually stop doing it?

S6: um, nn, not really

S3: no uhuh, because the

S6: i i haven't read any further

S3: because the actual, time they do it doesn't matter, it's it's whether there's an actual stage they come to and that's a very important principle of child language

S6: i haven't reached that point

S3: that it's the sequences that are important not actual, chronological age

S6: i haven't reached that part i'm still reading it, but i think she says well, they they still until at this age they mix and, i didn't continue, but, i_ here i would like just to finish to to talk about my own study that i intend to do exactly that, and, will analyze the speech of, English Spanish bilingual speakers and also have a look at these aspects to see if it's really, universal it's cross-linguistic or not, and for me this is what's interesting for that. i i'm g- gonna try to replicate Lanza's um, um study and see... mainly in these two (axes) um, um pragmatic and syntactic um

S3: have you have you read Andrew Radford's account of language acquisition, of English because he [S6: yes] talks about INFL the emergence of INFL [S6: uhuh] that's the basic

S6: yeah Radford, was, my very first contact with [S3: right] uh [S3: okay,] with, with, Chomsky's theories [S3: right] (with_ through) Radford and that's what we

{END OF TRANSCRIPT}

APPENDIX E: SAMPLE ACTIVITIES

E1:

AWARENESS RAISING ACTIVITY: DISCOURSE MARKERS

(designed by the researcher)

Aim: To raise awareness of students about types and functions of discourse markers explicitly.

Level: From intermediate to advanced students.

Skills: Speaking & Listening

Duration: 45 minutes.

Materials:

A Handout (see below)

A computer with internet connection

Description:

Teacher explains discourse markers and their functions (see teacher resource below)

Students are given information about what discourse markers are and their functions and they practice on finding out the functions of the discourse markers in a given transcript.

Procedure:

As a warm up, teacher writes on board the following words “*Oh, Well, You know, Like, Kind of, So*” and asks students what they can have in common.

After question-answer, teacher explains that they are called discourse markers and by using teacher resource below talks about what discourse markers. At this step, teacher can do question and answer session about why they are called discourse marker sor why do we use them to elicit students’ knowledge about them.

Teacher distributes a handout including a transcript taken from a native speaker’s student presentation (see handout) or optionally, s/he can have another transcript taken from dialogue of native speakers or else.

Pair-work: Student, in pairs, first tries to find out the utterances that can be discourse markers and then name their functions in the discourse according to the functional categories.

As a whole class discussion, teacher guides students to discuss which functions each pair have found and whether they are similar with the others.

Teacher Resource:

Definition of discourse marker: A [particle](#) (such as *oh*, *well*, *now*, and *you know*) that is used in [conversation](#) to make [discourse](#) more [coherent](#) but that generally adds little to the [paraphrasable meaning](#) of an [utterance](#).

In most cases, discourse markers are *syntactically independent*: that is, removing a marker from a sentence still leaves the sentence structure intact. Discourse markers are more common in informal [speech](#) than in most forms of [writing](#). (<http://grammar.about.com/od/d/g/discoursemarkerterm.htm>)

Functions of discourse markers (Fung and Carter, 2007)

Interpersonal	Referential	Structural	Cognitive
Marking shared knowledge: <i>See, you see, you know, listen</i>	Cause: <i>Because, cos</i> Contrast: <i>But, Now, OK/okay, and, yet, however, nevertheless</i>	Opening and closing of topics: <i>right/alright, well, let's start, let's discuss, let me conclude the discussion</i>	Denoting thinking process: <i>Well, I think, I see, and</i> Reformulation/Self-correction: <i>I mean, that is, in other words, what I mean is, to put it in another way</i>
Indicating attitudes: <i>well, really, I think, obviously, absolutely, basically, actually, exactly, sort of, kind of, like, to be frank, to be</i>	Coordination: <i>And</i> Disjunction: <i>Or</i> Consequence: <i>So</i> Digression: <i>Anyway</i>	Sequence: <i>First, firstly, second, secondly, next, then, finally</i>	Elaboration: <i>Like, I mean</i>

<i>honest, just, oh</i>	Comparison:	Topic shifts: <i>So, now, well, and what about, how</i>	Hesitation: <i>Well, sort of</i>
Showing responses:	<i>Likewise, similarly</i>	Assessment of the listener's knowledge about the utterances:	
<i>OK/okay, oh, right/alright, yeah, yes, I see, great, oh great, sure, yeah</i>		Summarizing opinions: <i>So</i>	<i>You know</i>
		Continuation of topics: <i>Yeah, and, cos, so</i>	

HANDOUT:

An extract from MICASE (Michigan Corpus of Academic Spoken English, English Language Institute, University of Michigan) From <http://micase.elicorpora.info/sound-files-online>, it is also possible to listen the student presentation titled Bilingualism Student Presentations.

S1: okay, now basically this is uh, a study on um, i wanted to look at um, the bilingual population here at the University of Michigan, and i started out thinking i would just look at, whatever people i could find who were bilingual, and then i realized that it would be difficult to kind of, find anything out about it if i was just looking at whatever languages i came up with. so i ended up concentrating on two different languages, and um, i also concentrated on the easiest samples i could find, so i have a Spanish um, English bilingual sample, taken from a class actually of self identifying U-S Latinos. so that's something to consider, and then i have um, a sample from, a Hindi class, uh two Hindi classes actually here, which are mostly bilingual um heritage speakers of Hindi, or another Indian language. um, and, uh basically, no one other language is highly represented besides Hindi but i'll talk about the littl- that a little bit in a minute. um the things that i was really looking for were, uh, had to do with, theoretical affiliation, with um a language community, which might serve as a community using these uh measures, basically,

S2: Helen

S1: yeah?

S2: can you put the window, down?

S1: the which?

S2: the window, down?

S1: the window down? <NOISE DISRUPTION> see if i can get those um... <NOISE DISRUPTION><SS: LAUGH> um, that helps a little [SU-f: (xx)] okay, um okay the first thing i was looking at is theoretical affiliation like do you think that you will be basically affiliated with this language community in this way in the future (that's what) (xx) um so first i ask, i just sort of_ these are just theoretical right though they have nothing to do with what we are doing now and i'll, talk a little bit about one other thing they ask

that's about right now, in a minute but these're basically, the main things i was looking at. and i also looked at the importance of maintaining or increasing their proficiency in language A, um, for the various reasons. and, what i ended up doing in terms of figuring out what these meant and how they correlated was, looking at um, basically what i did for these scores or these answers, is i gave them two, points if they, said yes one point if they say maybe and zero points if they say no so i have an aggregate score for that, and i have um an average, score for this one. um, depending on various things and i also i'll show you a little bit about how those broke down, um... there's a lot of different things you can consider in terms of what will influence, you know how someone, perceives their own identity and how important it is for them to maintain their other language. um basically with students at the U-of-M, we're looking at people who, um, are basically social- socially mobile, so um, i sort of went in assuming that i would find people who would be more tending to be agents of language shift, rather than language retention, um in the sense that they, most of the students um, like in the in the Spanish, bilingual, Spanish English bilingual sample, three out of eleven, had grown up in a Latino neighborhood or, i_ whatever, i i measured that just by saying, you know, do you h- did you have neighbors who were spoke Spanish did you have um classmates who spoke Spanish? and only three of them answered yes to either of those. so basically most of them did not grow up in a bilingual community, except for that of their parents, and i've done qualitative interviews i haven't, finished evaluating all of that yet but basically, for the qualitative, work it seems like most of the students, did indeed grow up in um an integrated community, and they often say well i spoke Spanish or Hindi with my parents friends or their childre- their par- my parent's friends' children, but i didn't, you know have friends at school who spoke it or whatever. um another thing to consider of course is the different language groups, um, Hindi obviously, uh, is not, as widespread in the United States as Spanish is, um so in terms of usefulness or, accessibility in the United States it's a little different. so i'm just gonna look at some differences between the two groups, for a moment here, um, this is the, answer to th- the second question which is about the importance of, um, th- these various reasons for maintaining or increasing, fluency in language A (alright,) um, and i've got this little, graph at the bottom, which basically tells you that, i- i've sort of labeled this so, um, basically for the Hindi class, seventy-five percent answered_ had an average of this or higher, the seventy-fifth percentile, half of them had this or higher, twenty-five percent, so you can see that the Hindi class has a much, um, lower range basically, of answers okay, and you can basically, see that that's from these two, numbers here, i circled them here, um, okay, um so for professional and academic reasons, Hindis, were rated not that much more- much lower than the Hindi speakers- than the Spanish speakers. um, and that was pretty much across the board, and, we can think of some obviously um reasons for that, uh... another interesting thing you can see there's some variability um i don't think that the rest of it's very significant though, um, the variability in the different ranges, um there is some variability though there is still like, this is relatively low compared to the other scores, for Spanish speakers, to the academics but it's still relatively high, this was just- a rating of four just means important, anyway so, and then, just to look at the other, thing i was interested in measuring, this is the questions on theoretical affiliation, and i broke these down into proportions so this is like basically like thirty-six percent, (xx) fifty-five percent. um, so for this- for the Spanish um, the Spanish English bilingual sample, or the Spanish class i should just say um, basically, there's a lot more, um, higher ratings in the in the yes category like, a lot more people who said yes to um, to everything, and there's also very few people who said no to anything. you'll see here most people said at least maybe, you know, or yes. um and that again might have something to do with the also the sample is taken from a U-S Latino_ self-identifying U-S Latino class, and so students who are in this class presumably are already relatively invested in being a U-S Latino in the sense that they self-identify um, but at the same time, um the Hindi speakers, um, no one said no for example, no they would not marry a Hindi speaker, but um, you know two people i believe that represents two people who said they would never marry someone who was a Spanish English bilingual. basically. okay. um, and you can see again there's, a higher distribution for the Spanish speakers than the Hindi speakers for the aggregate score, like all these together, um, (let's) see and the other interesting things, th- only one person i think said yes they would live in a country where Hindi was spoken basically India right, um, whereas, um several people said yes they would, (live, in) live in a Spanish speaking country, definitely and um, let's see, the other interesting thing is that hardly anybody said that, they would not raise their children bilingually, (okay) in any of the samples there's two people total, in both samples that said that they, would would not raise their children bilingually. or maybe actually they didn't say would not they said maybe i won't. you know <LAUGH> and one of them like even wrote a comment in like i'm not sure if i'm proficient enough to bring my children up bilingually <LAUGH> like i want to but, so basically you can see by just, looking at that that, basically for all of the students sampled they are all pretty invested in, thinking that bilingualism is a good thing, you know, over all, something they want to pass on to their children, in the interviews that comes out as well... um, so i was interested also in how these correlate i was sayi- thinking you know m- well maybe some people are interested in being part of this community the bilingual community, but they're not really interested in, um, having, uh, high, their (own,) high proficiency (in the) language, so there is a high, (um,) there is a high correlation between these two things so, this is the average, this is the average score of like the five point rating thing and this is the aggregate score of like yes maybe no, do you want to do this do you want to do that, um, and there's like a high correlation i think it's point, seven two.

E2:**AWARENESS RAISING AND PRACTICE ON DISCOURSE MARKERS**

(designed by the researcher)

Aim: To raise awareness of the students about the use of discourse markers and their functions and to practice them.

Level: From intermediate to advanced.

Duration:

Duration of out of class part cannot be given.

Duration for on class activity: it depends on the student population.

Materials:

A computer with internet connection

Audio-video recorder of the students (students' cell phones with these qualities may help)

Skills: Speaking and Listening

Description: This activity includes language learners' receiving and producing discourse markers in their spoken discourse and increase in awareness of using discourse markers functionally.

Procedures:

Before class, (probably at the previous course) teacher gives a pair-homework to the students which is about making an interview with the partner by asking three questions about their city. The questions are "*What is(the city) really like? What' the best thing in (the city)? What is the worst thing in (the city)?*" and this interview should be audio and/or video recorded and should be transferred to the computer before the class.

In the beginning, as a warm up, teacher can explain what shall be done during the course. Teacher plays the videos of the students' interviews. Two or three pairs (teacher can decide the number) voluntarily plays their videos so other students also can watch the videos.

At the next step, teacher plays the video podcasts named as London Native Speaker Interviews 1,2,3 (<http://www.youtube.com/watch?v=Uvh6ra1gbjc>). Then, as a whole class discussion, the question "*what kind of differences are there between their interviews and native speaker interviews?*" is discussed.

Teacher draws students' attention to the items "so, well, right, really, etc. and explains their significance (teacher resource above is used) in spoken discourse and their use by native speakers.

In pairs, students again do the same interviews by trying to use the discourse markers that they have learned.

E3:**Discourse markers *Well* and *Oh***

(by Kent Lee, University of Illinois at Urbana-Champaign, United States,
<http://exchanges.state.gov/media/oelp/teaching-pragmatics/lee-well.pdf>)

Level: Lower to upper-intermediate ESL/EFL, such as at an intensive English program

Time: 30-35 minutes

Resources: Sample sentences and situations for role play

Goal: Using the discourse markers *well* and *oh* for smoother discourse flow.

Description of the Activity

The teacher begins with a mini-role play with some students by asking each student to request a favor from the teacher (“Could I borrow your car?”, “Can I borrow a dollar?”, “Could you take me to the airport next weekend?”). The teacher answers quickly without pauses or discourse markers, thus sounding abrupt, such as “No, I think I need it tonight” or “I’m afraid I don’t have any money”, “I’m going to be out of town next weekend”). The teacher repeats the situation with another student but provides smoother responses, with delays and realizations marked by *well* and *oh* (“Well, I think I need them to study tonight, “Oh/Well... I’m afraid I don’t have any Money”, “Well... oh, I’m going to be out of town next weekend”). The process can be repeated with other situations, such as making difficult requests (“Uh, could I borrow your car?”), deflecting complements, and pauses (such as those as elicited by difficult or face-threatening questions like “How old are you?”).

The teacher solicits the students’ impressions of the second set of answers compared with the first. The students should recognize the second set as more polite and “smoother.” The teacher queries the students as to why they sound better, and the purposes and functions of the words *well* (unexpected response) and *oh* (realization).

The students may not ascertain these 2 abstract discourse functions other than filling pauses and delays, but the teacher can use students' guesses to build up to an explanation of their discourse functions described above *B well* indicating negative or unexpected responses, and *oh* indicating sudden realization necessitating a shift to another topic or to an unexpected response. With sample dialogues, the teacher can illustrate other related functions of *well* (turn-taking, topic shift or resumption) and *oh* (repair, clarification), elicits students' impressions as above, and elucidates their functions. Then students can be provided with similar situations for them to practice with each other in small groups or pairs. Situations may require functions such as issuing and declining requests and compliments, making difficult requests, turn-taking, competitive turn-taking, answering difficult questions requiring repair and pauses for planning, topic shift, apologies, responses involving "realizations", topic shift, topic resumption, repair, and clarification.

Procedure

1. Teacher-student mini-role play [10 minutes]

- a. The teacher asks several students to request a favor from the teacher.
- b. The teacher refuses with slightly abrupt answers lacking discourse markers or other delays.
- c. Next the teacher does so with refusals marked by delays and discourse markers.
- d. The process is repeated with a few other linguistic functions: difficult requests, deflecting compliments, and apologies.

2. Discussion. [5-10 minutes]

- a. The teacher solicits students' impressions of the two sets of answers, including: why the second version sounds better, and the purposes and functions of *well* (delay, and unexpected response) and *oh* (realization).
- b. Drawing from students' responses and the role-play situations, the teacher elucidates the functions of the discourse markers.

3. Student role play [10-15 minutes]

- a. Students are given more complex situations to role play conversations in groups of 2-3.
- b. The teacher can circulate to help or coach students in their conversational role-plays.
- c. Optionally, the teacher may select a few groups to present their role-plays before the whole class.

Teacher Resource

Conversation excerpts with “like” and “oh”

A: How much education do you think a person needs to get a good job?

B: Oh, definitely a bachelor’s degree.

A: Well, I think even more than that. At least a master’s degree.

A: She can listen and tell you not only the composer but the name of the piece.

B: Well, that’s no big deal.

A: Who wants to know?

B: Well, I want to know.

A: Can I borrow your car?

B: Well...my wife needs to use it tonight.

A. ...Well, like I was saying, I think the only difference between our neighborhoods might be the better trash collection in our neighborhood.

A. I think that law was passed in 1976. Oh, maybe it was 1978, I don’t remember for sure.

A. How can I get a grant for that?

B: Oh, I didn’t realize they gave grants. I’m not the one to ask about that.

E4:**Several Activities for Using ‘OK’ in English**

(by Lindsay Clandfield, Macmillan Publishers Ltd 2004, Downloaded from www.onestopenglish.com?)

The following activities are designed to help your students become aware of the importance of ‘OK’ and to begin to use it in conversations.

Language Information box:**How do people use ‘OK’ in conversation?**

The most basic use of ‘OK’ is to accept a suggestion, request, offer or information.

Mother: I don’t want you to come home later than 11 o’clock.

Daughter: OK.

You can use ‘OK’ to show that you accept someone’s response but you have something else to say about the situation.

Football player: I can’t run very fast. My leg hurts.

Coach: OK, you don’t have to come to practice today.

Another use for ‘OK’ is to join two topics or stages of a talk.

Teacher: We’re studying a new grammar point, and I gave some homework.

Can you check it with them?

Substitute teacher: OK. Is there anything else I should do?

‘OK’ is often used by professors or teachers to move on to another topic in the class. It is used to move on to the next item of business in meetings.

Finally, ‘OK’ is used to close a conversation, or when a conversation is finishing. It is often used with bye or see you.

Girlfriend: So will you come to my house tomorrow?

Boyfriend: Yes, I promise.

Girlfriend: OK, see you.

Boyfriend. OK, bye.

Girlfriend: Bye.

1. The Etymology of 'OK'

Aim: To find out about the origins of 'OK'.

Skills work: Reading and speaking

Divide the class into pairs A and B. Give each student A a copy of worksheet A and each student B a copy of worksheet B (below). Explain that they have to ask each other questions to find out the story of the word 'OK'.

Where does 'OK' come from?

'OK' is the most successful of all Americanisms. The word has spread to several other languages. However, its origins are not clear and there are several explanations as to where it came from. The study of where words come from is etymology. In this activity you will find out some of the etymology of the word 'OK'.

Work with a partner. One of you has worksheet A, the other has worksheet B. Ask questions to each other to find the missing information on the worksheet. The origin of 'OK' has been one of the most popular subjects in etymology. Here are some of the popular explanations of how the word was created.

1. 'OK' stands for O (zero) Killed. American soldiers used it in _____ when they came back from missions. Zero killed meant that no soldiers had died in that mission.
2. 'OK' comes from the Native American word okeh. Nobody knows what this original word okeh meant.
3. 'OK' is an abbreviation for _____ (the German for Colonel-in-Command) and was used by Baron Von Steuben, an American General in the Revolutionary War.
4. 'OK' comes from the French name for a port in Haiti called Aux Cayes. Aux Cayes was famous for its rum.
5. 'OK' stands for Orrin Kendall. Orrin Kendall was a _____ supplied to American troops during the American Civil War.
6. 'OK' was originally a 1860s telegraph term for Open Key (the same kind of key that you can find on a modern computer keyboard).
7. 'OK' stands for oll korrek, a misspelling of all correct. _____ was a terrible speller and wrote 'OK' for all correct on documents. Many etymologists say that this explanation is closest to the truth.

2. Are you listening 'OK'?

Aim: To recognise when and how often native speakers use okay in speech.

Skills work: Listening

Bring in a piece of authentic English listening text of informal conversation (from the radio, or a film in English. You can also use the Web to record audio texts.). Tell the students to listen out for the number of times someone says 'OK' (you must obviously choose a text which has 'OK' in it. ☺). Play the tape or video again and ask students to

write down the exchange in which 'OK' was used. Use this data to raise students' awareness of how 'OK' is being used and draw out the different ways it can be used. Make copies and distribute (or write on the board) the different uses of 'OK' from the language information box above to follow up. It is very useful to look at a learners' dictionary definition for 'OK' at this stage. Try the [Macmillan English Dictionary](#).

3. 'OK' in class

Aim: To recognise when and how often the teacher uses 'okay'.

Skills work: Listening

At the beginning of the week, ask the students to listen out for when you as a teacher use the words 'OK' (teachers use this discourse marker quite a lot!). You could ask someone to keep a record of how many times you use it during the classes that week. At the end of the week ask students to report back on your use of 'OK'. Use this information in a similar way to 2 above.

4. How do you say 'OK'?

Aim: To practise intonation and using different pitch.

Tell students that you want them to practise using their voices to express different emotions. Tell them that they are going to do this by using one of the most popular words in the English language, 'OK'. Write the following words on the board:

HAPPY, FRUSTRATED, CONFUSED, CONFIDENT, TIRED, SCARED

Demonstrate yourself by saying 'OK' in different ways and asking students to guess which emotion you are expressing. Then point to the different emotions and ask students to say 'OK' in the manner of that emotion.

5. Wrangling*

Aim: To practise intonation and using different pitch.

Skills work: Speaking

In a wrangling activity, students practice a two-line dialogue in pairs. The students say the lines to each other as an argument: they are allowed only to use the words in the dialogue but can use intonation and gesture to convince each other. The first to give up is the loser. You can start this activity by having two students volunteer to do this in front of the class.

Here are some sample dialogues to wrangle with, incorporating ‘OK’.

OK, then finish it later.

But I’m very busy.

I didn’t do it.

OK, prove it.

I think this is stupid.

OK, but we’re learning.

Would you like anything else?

No, it’s OK.

* The idea for this wrangling activity is adapted from Penny Ur’s book *Five Minute Activities* (CUP, 1992)

6. OK or No Way

Aim: To practise using discourse markers to respond to suggestions

Skills work: Speaking

Write on the board the following.

OK.

No way.

Clarify the pronunciation and meaning of each of these expressions (“no way” is an informal and direct way of refusing a suggestion). Then explain that you are going to make several different suggestions to the class. They must respond with either OK or no way, depending on whether or not they want to accept the suggestion. Make a few suggestions, first to the class as a whole (who must respond chorally) then to

individual students. When they get the idea, ask them to continue in pairs. They can further elaborate their conversations by giving reasons for accepting or refusing.

Sample suggestions:

Would you like more homework for tonight?

Would you like no more tests this year?

Would you like \$100?

Can you give me \$100?

Can I have your book?

Can I borrow your car?

Can I take a picture of you?

Note: You could combine this aspect of using OK with any lesson you do on the function of making suggestions, or teach different ways of making suggestions (Would you like..., How about ...+ ing, Why don't we...).

7. OK, but...Generalisation game

Aim: To practise disagreeing using discourse markers.

Skills work: Speaking

While OK is often used to accept a suggestion, curiously enough it also serves the purpose of refusing one. In this case, the OK is said to acknowledge the other person's contribution and is followed by a counter suggestion or contribution (often prefaced with a but).

Ask the students to work in pairs. Give each pair a copy of the card below and ask them to read the instructions. Explain that students must choose only one of the three words to begin the activity. They must then try to keep their conversation going for at least two minutes. When two minutes are up, ask students to swap roles and write up three new words for students to choose from.

Sample words you could use for this game:

lawyers, famous football players, grandparents, doctors,

Americans, teenagers, people who work in public administration,
teachers, police officers.

The GENERALISING GAME - Instructions

Work with a partner for this activity. One of you is A, the other is B.

Your teacher is going to write three words on the board. Student A chooses one of them.

Student A must make a generalisation or comment about the word on the board.

Student B must disagree or present an exception to this rule.

Student A can disagree with Student B, or go onto the next word.

Useful language to make a generalisation:

Generally speaking...

Most of the time...

By and large...

Useful language to start disagreeing:

OK, but what about...

OK, but think of...

OK, but there are exceptions of course...

8. OK, where was I?

Aim: To practise using discourse markers to return to a topic in conversation

Skills work: Speaking

This is a variation on a favourite English Language Teaching activity. Prepare a story or joke in advance of the class. Make sure it is something you can remember well and that you can recite. Explain to the students that you are going to tell them the story, but that they must interrupt you and ask questions. Give an example to help them get started. When students interrupt, accept only to answer their question if it is formed correctly. When you have answered the question, come back to the point you left off by saying one of the following phrases:

OK, where were we? Oh yes...

OK, so...

OK, as I was saying...

(You can make your own of phrases that you would use under those circumstances).

Once you have (finally) finished your story, write the language you were focussing on (the phrases above) on the board and draw students' attention to them. Students can do the same activity in pairs.

9. OK, bye!

Aim: To practise using discourse markers to close a conversation.

Skills work: Speaking

For this activity, divide the class into As and Bs.. Give the As a couple of minutes to prepare an important story or series of instructions which they are going to tell B.

While the As are preparing explain to the Bs that they are going to listen to A, but that they are desperately anxious to get away. B must try to stop A's conversation by occasionally interrupting with the words OK.

Put the students together in pairs, A and B. Ask A to begin their story while B tries to get away from the conversation. Circulate and check. You could get a strong pair to perform this in front of the class.

OK, anything else? A closing word

An overuse of the discourse marker OK might seem sloppy or informal to some speakers. The ultimate aim of many of these activities is to allow students to incorporate OK into their natural speech. Used correctly, a discourse marker such as OK can make one's English sound a lot smoother. It is also better than 1) going silent in the middle of a conversation or 2) using a discourse marker in your own language when speaking English, especially if this is likely to be misunderstood by the other speaker.