DISCOURSE MARKERS IN NATIVE AND NON-NATIVE SPOKEN ENGLISH: A CORPUS-BASED COMPARISON OF TURKISH AND BRITISH UNIVERSITY STUDENTS' EMPLOYMENT OF DISCOURSE MARKERS IN INFORMAL INTERVIEWS

Doktora Tezi

Yusuf ÖZTÜRK

Eskişehir 2018

DISCOURSE MARKERS IN NATIVE AND NON-NATIVE SPOKEN ENGLISH: A CORPUS-BASED COMPARISON OF TURKISH AND BRITISH UNIVERSITY STUDENTS' EMPLOYMENT OF DISCOURSE MARKERS IN INFORMAL INTERVIEWS

Yusuf ÖZTÜRK

PhD THESIS

Department of Foreign Language Education
PhD in English Language Teaching
Supervisor: Prof. Dr. Gül DURMUŞOĞLU KÖSE

Eskişehir

Anadolu University

Graduate School of Educational Sciences

May 2018

This PhD thesis has been funded by the Scientific Research Unit of Anadolu University.

Project No: 1305E093

APPROVAL OF JURY AND THE INSTITUTION

(JÜRİ VE ENSTİTÜ ONAYI)

Yusuf ÖZTÜRK'ün "Discourse Markers in Native and Non-Native Spoken English: A Corpus- Based Comparison of Turkish and British University Students' Employment of Discourse Markers in Informal Interviews" başlıklı tezi 28.05.2018 tarihinde aşağıdaki jüri tarafından değerlendirilerek "Anadolu Üniversitesi Lisansüstü Eğitim-Öğretim ve Sınav Yönetmeliği"nin ilgili maddeleri uyarınca Yabancı Diller Eğitimi Anabilim Dalı İngilizce Öğretmenliği Doktora Programında, Doktora tezi olarak kabul edilmiştir.

	Unvanı-Adı Soyadı	İmza
Üye (Tez Danışmanı)	: Prof. Dr. Gül DURMUŞOĞLU KÖSE	
Üye	: Prof. Dr. Ümit Deniz TURAN	
Üye	: Prof. Dr. Gölge SEFEROĞLU	
Üye	: Doç. Dr. İlknur SAVAŞKAN	
Üye	: Doç. Dr. Özgür YILDIRIM	

Prof.Dr. Handan DEVECİ
Anadolu Üniversitesi
Eğitim Bilimleri Enstitüsü Müdürü

ABSTRACT

DISCOURSE MARKERS IN NATIVE AND NON-NATIVE SPOKEN ENGLISH: A CORPUS-BASED COMPARISON OF TURKISH AND BRITISH UNIVERSITY STUDENTS' EMPLOYMENT OF DISCOURSE MARKERS IN INFORMAL INTERVIEWS

Yusuf ÖZTÜRK

Department of Foreign Language Education
PhD Programme in English Language Teaching
Anadolu University, Graduate School of Educational Sciences, May 2018

Supervisor: Prof. Dr. Gül DURMUŞOĞLU KÖSE

Discourse markers serve a variety of functions in spoken discourse and are valuable for being pragmatically competent in a language. However, non-native speakers are reported to differ from native speakers in their use of discourse markers. Moreover, the literature on the use of discourse markers by Turkish EFL learners has been limited, and the research available is restricted to mostly planned and monologic speech. In this regard, this study aims to investigate the use of five most common discourse markers, namely *so*, *like*, *you know*, *I mean* and *well*, by Turkish and British university students, and examine the relationship between the Turkish students' discourse marker use and speech fluency. A corpus of informal interviews was compiled of interviews with 50 Turkish ELT students, and a native speaker corpus created at a UK university was obtained. The results showed that Turkish students used the discourse markers in textual functions in a higher proportion and interpersonal/interactional functions in a lower proportion compared to the British students. The first-year Turkish students employed the markers significantly more frequently than their fourth-year peers. Moreover, there was a moderate, and significant relationship between discourse marker use and their utterance and perceived fluency. These results were discussed in the light of the relevant literature, and various implications and suggestions were offered for teaching and further research.

Keywords: Discourse markers, Turkish students, British students, Informal interviews, Speech fluency.

ÖZET

ANADİL VE YABANCI DİL KONUŞURU SÖZLÜ İNGİLİZCESİNDE SÖYLEM BELİRLEYİCİLERİ: TÜRK VE İNGİLİZ ÜNİVERSİTE ÖĞRENCİLERİNİN İNFORMAL GÖRÜŞMELERDE SÖYLEM BELİRLEYİCİLERİ KULLANIMININ DERLEM TEMELLİ BİR KARŞILAŞTIRMASI

Yusuf ÖZTÜRK

Yabancı Diller Eğitimi Anabilim Dalı
İngilizce Öğretmenliği Doktora Programı
Anadolu Üniversitesi, Eğitim Bilimleri Enstitüsü, Mayıs 2018

Danışman: Prof. Dr. Gül DURMUŞOĞLU KÖSE

Söylem belirleyicileri sözlü söylem içerisinde bir dizi işleve sahiptir ve bir dilde edimbilimsel olarak yetkin olmak için önemlidir. Ancak, yabancı dil konuşurlarının söylem belirleyicileri kullanımında anadil konuşurlarından farklılaştığı bulgulanmıştır. Ayrıca, Türk İngilizce öğrenenlerin kullanımıyla ilgili alanyazın sınırlıdır ve mevcut araştırmalar planlı ve monolojik konuşmaya dayalıdır. Bu anlamda, bu çalışma (a) Türk ve İngiliz üniversite öğrencilerinin beş söylem belirleyicilerini (yani *so*, *like*, *you know*, *I mean* ve *well*) kullanımını araştırmayı ve (b) söylem belirleyicileri kullanımı ile konuşma akıcılığı arasındaki ilişkiyi incelemeyi amaçlamaktadır. 50 Türk İngilizce öğretmenliği öğrencisi ile informal görüşmeler yapılmış ve bir İngiliz üniversitesinde oluşturulmuş anadil konuşuru derlemi edinilmiştir. Sonuçlar, Türk öğrencilerin bu belirleyicileri daha az sıklıkta kullandıklarını ve İngiliz öğrencilere göre bunların metinsel işlevlerini daha çok, kişilerarası/etkileşimsel işlevlerini ise daha az kullandıklarını göstermiştir. Türk 1. sınıf öğrencileri, dördüncü sınıf öğrencilerine göre, incelenen belirleyicileri daha sık kullanmışlardır. Ayrıca, Türk öğrencilerin bu belirleyicileri kullanımı ile konuşma akıcılıkları arasında orta düzey anlamlı bir ilişki bulunmuştur. Bu sonuçlar, ilgili alanyazın ışığında tartışılmış, öğretim ve gelecek araştırmalara yönelik çıkarımlar çalışmanın sonunda sunulmuştur.

Anahtar Sözcükler: Söylem belirleyicileri, Türk öğrenciler, İngiliz öğrenciler, İnformal görüşmeler, Konuşmada akıcılık.

ACKNOWLEDGEMENTS

Pursuing a PhD degree was a long, exhausting and difficult process in which I encountered a number of challenges that I handled with the help, guidance and encouragement of many other people around me.

First of all, I am sincerely grateful to my thesis supervisor, Prof. Dr. Gül DURMUŞOĞLU KÖSE, for always being there whenever I needed help. Her conscientious feedback and constructive criticism were invaluable for me. I believe this thesis would not have been completed without her guidance.

I would like to thank my thesis committee members, Prof. Dr. Ümit Deniz TURAN, Prof. Dr. Gölge SEFEROĞLU, Assoc. Prof. Dr. İlknur SAVAŞKAN and Assoc. Prof. Dr. Özgür YILDIRIM for their feedback and contributions to this thesis.

I want to thank all my teachers, colleagues and the head of the ELT department at Anadolu University, Prof. Dr. Ümit Deniz TURAN. I feel privileged to have studied and worked at this special department for over seven years during my MA and PhD studies.

I owe many thanks to the students whose names I could not mention throughout this study, but who contributed to this work as interviewees.

I would like to thank my dear friends, Res. Asst. Tuncay KARALIK, Res. Asst. Musa TÖMEN and Res. Asst. Samet TAŞÇI, for their help in many stages of this work. I am also grateful to Ian FITZMORRIS and Lyndsey REED for their contributions as native English speakers.

I must thank my whole family for their support and understanding in my entire life. Above all, I would like to express my gratitude to my late grandmother, Azize BAHADIR, who passed away in late 2017, for taking care of me when I was a little boy, and when I just started my undergraduate education away from home. Thank you for all the things you did for me!

I am indebted to my parents, Ercüment and Zehra ÖZTÜRK, for always supporting me in every step of my life. I have always felt their prayers whenever I faced a difficulty.

I also owe a gratitude to my aunt, Prof. Dr. Füsun ÖZTÜRK KUTER, for encouraging me to pursue an academic career when I just completed my BA. I learned

many things from her about life and the academia, which was of great importance to me.

Finally, and most importantly, I would like to thank my beloved wife, Esra ÖZTÜRK because writing this PhD thesis would not be possible without her love, support and understanding. There is one last person to whom I am also grateful; my son, Ali Bera, and I owe him an apology as well. Sorry for the long hours away while you wanted to play with me, but I promise I will make it up to you.

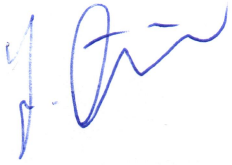
Yusuf ÖZTÜRK

Eskişehir 2018

08/05/2018

STATEMENT OF COMPLIANCE WITH THE ETHICAL PRINCIPLES AND RULES

I hereby truthfully declare that this thesis is an original work prepared by me; that I have behaved in accordance with the scientific ethical principles and rules throughout the stages of preparation, data collection, analysis and presentation of my work; that I have cited the sources of all the data and information that could be obtained within the scope of this study, and included these sources in the references section; and that this study has been scanned for plagiarism with "scientific plagiarism detection program" used by Anadolu University, and that "it does not have any plagiarism" whatsoever. I also declare that, if a case contrary to my declaration is detected in my work at any time, I hereby express my consent to all the ethical and legal consequences that are involved.



Yusuf ÖZTÜRK

TABLE OF CONTENTS

Page

Page

Page

Page

LIST OF TABLES

Page

Page

LIST OF FIGURES

Page

LIST OF ABBREVIATIONS

BASE	: British Academic Spoken English
CANCODE	: Cambridge and Nottingham Corpus of Discourse in English
DMs	: Discourse Markers
EFL	: English as a Foreign Language
ELT	: English Language Teaching
ESL	: English as a Second Language
MICASE	: Michigan Corpus of Academic Spoken English
L1	: First Language
L2	: Second Language
LINDSEI	: Louvain International Database of Spoken English Interlanguage
LOCNEC	: Louvain Corpus of Native English Conversations

1. INTRODUCTION

1.1. Background to the Study

Pragmatic competence is a crucial component of communicative competence (Bachman, 1990, p. 87; Bachman and Palmer, 1996, p. 69). It is the ability to communicate one's intended message along with its nuances in a particular context and interpret his/her interlocutor's message in the way it is intended (Fraser, 2010, p. 15). A key element that is argued to play a significant role in the pragmatic competence and spoken interaction of a speaker is discourse markers (Müller, 2005, p. 1; Mullan, 2010, p. 9). Discourse markers are lexical items that individuals use to create textual coherence and express their feelings and views (Carter and McCarthy, 2006, p. 208), and are argued to be pragmatically indispensable in spoken discourse (Mei, 2012, p. 2). They are words or phrases such as *so* and *well* that individuals employ to connect, organise and manage what they say or to express their attitudes:

(1a) A: I like him. B: *So* you think you'll ask him out then.
(1b) A: Did you like it? B: *Well*, not really. (Fraser, 1990, p.383)

With the emergence of many spoken corpora, such as The Michigan Corpus of Academic Spoken English (MICASE) and the British Academic Spoken English (BASE), there has been an expanding literature addressing spoken English, an area where the use of discourse markers has been a popular research topic. In the literature, native speakers' use of discourse markers has been extensively investigated in many studies including Schourup (1985), Schiffrin (1987), Fraser (1990, 1999), Jucker (1993), Lenk (1995), Aijmer (2002) and Carter and McCarthy (2006), which provided valuable results. As for non-native speakers, their use of discourse markers attracted little, though increasing, attention in the recent decade. However, it is also important for non-native speakers who are expected to interpret native speakers' use of discourse markers and to use them in appropriate contexts and situations (Huang, 2011, p. 11). Discourse markers form useful contextual coordinates for both native and non-native speakers of English to structure and organise speech as they perform important textual and interpersonal functions (Fung and Carter, 2007, p. 411).

Discourse markers arose from the close study of spoken discourse and is predominantly a feature of spoken rather than written discourse (Brinton, 1996, p. 33; Müller, 2005, 27; Mullan, 2010, p. 44). According to Schiffrin (1987, p. 31), the term refers to "sequentially dependent elements which bracket units of talk" . For Fung and Carter (2007, p. 411), these units are intra-sentential and supra-sentential linguistic elements having non-propositional and connective functions at discourse level. Lenk (1998, p. 52) defines discourse markers as "short lexical items" that are employed "to signal for the hearer how the speaker intends the present contribution to be related to preceding and/or following parts of the discourse". These definitions indicate that discourse markers ensure coherence within and among utterances as well as turns in discourse.

Discourse markers, like "*well*", "*I mean*" and "*you know*", are highly frequent in conversation. "*You know*" and "*I mean*" are reported to be the most frequent two-word combinations in the Cambridge and Nottingham Corpus of Discourse in English (CANCODE), a spoken corpus of British English. Though each of these discourse markers has a simple lexical meaning, they also have a function that they fulfil in discourse, and it is argued that this function contributes most to their high frequency (O'Keeffe, McCarthy, and Carter, 2007, p. 172). Furthermore, one of the most important properties of discourse markers is their multi-functionality, which means that they usually have more than one function in discourse. For example, *so* mostly marks a result or consequence, but it also has a function of marking a potential turn-transition (Schiffrin, 1987, p. 223). Therefore, the specific function of a discourse marker is determined by the context in which it is used (Aijmer, 2004, p. 177), which can make its acquisition difficult for learners/non-native speakers. At this point, there are two aspects that seem to have a role in learners' acquisition of discourse markers: informality of the learning context, and the functional levels of the discourse markers concerned. Firstly, since discourse marker use is regarded as being related to informal contexts, competence in this area can be acquired in such contexts (Müller, 2005, p. 47), which poses a difficulty for EFL learners if they are exposed to English only in a formal school environment. Secondly, discourse markers that functionally operate at the ideational or

textual levels are to be expected earlier than those that operate at other planes/levels such as interactional and interpersonal, because the former are overtly taught (Hays, 1992, p. 24). This also poses a challenge for learners because they presumably have difficulty in using discourse markers in non-textual functions.

1.2. Statement of the Problem

Many studies in the literature reported that non-native speakers of English differ from native speakers in their use of discourse markers (e.g. Romero Trillo, 2002; Fung and Carter, 2007; Buysse, 2010). In these studies, the differences include non-native speakers' preferences for different discourse markers (Romero Trillo, 2002), their restricted use with limited functions (Fung and Carter, 2007), and overuse or underuse of certain discourse markers (Buysse, 2010). Aijmer (2002, p. 3) argues that non-native speakers' incorrect use or underuse of discourse markers may result in misunderstandings. Moreover, they can be perceived as dysfluent when they demonstrate non-target-like use of discourse markers (Hellermann and Vergun, 2007, p. 160). Consequently, discourse marker use, which is performed by native speakers effortlessly, is also necessary for learners to express themselves in a fluent and confident way in that language (Sankoff et al., 1997, p. 214). The importance of being competent in using discourse markers particularly for non-native speakers is highlighted by Svartvik as follows:

> [I]f 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" (1980, p. 171).

Svartvik (1980) points out that learners' underuse or misuse of discourse markers could be more problematic than their grammatical mistakes in conversation. It thus seems self-evident that mastering the use of discourse markers should be equally important as the acquisition of grammatical competence. In this regard, learners' communicative competence would be greatly impaired if they did not understand the meaning of discourse markers (Wierzbicka, 1976, cited in Mei, 2012, p. 3). In other

words, discourse markers have an empowering function, in that their absence leaves individuals potentially disempowered in conversational interaction and puts them at risk of becoming a second-class participant (O'Keefe, McCathy and Carter, 2007, p. 39).

Given the importance of discourse markers provided above, previous studies on the use of discourse markers in conversation focused on learners from a variety of L1 backgrounds (e.g. Chinese, Japanese, Flemish, etc.). However, the literature on Turkish learners of English has been extremely limited in terms of both quantity and scope. For instance, Aşık and Cephe (2013) investigated the use of discourse markers in the spoken English of Turkish EFL (English as a Foreign Language) learners (i.e. ELT, or English Language Teaching, teacher candidates), but their examination included in-class student presentations, or in other words planned speech. Yet, discourse markers are more likely to be used in dialogic genres or in situations where more than one speaker is involved (Huang, 2011, p. 348), which is not actually the case in planned speech. Therefore, it can be argued that there is a need for further studies to explore Turkish learners' use of discourse markers and compare their use to that of native speakers of English in more spontaneous or natural conversation. The fact that discourse markers are not used by L2 learners to the degree that native speakers would use them (Hellerman and Vergun, 2007, p. 162) may be partly because discourse markers are not the focus of classroom language teaching and it is difficult for non-native speakers to acquire the discourse markers with functions that are more implicit to them (Nikula, 1993, p. 136). Furthermore, non-native-like use of such pragmatic items may lead to being marked as separate from the target speech community and inhibit chances for continued meaningful interaction (Boxer, 1993).

In addition to being highly frequent and useful in conversation, discourse markers are also argued to contribute to speakers' speech fluency through the functions they fulfil in discourse (Olynyk et al., 1987, p. 121). If, as is argued, they play a role in speakers' speech fluency and how fluent they sound to hearers, then this potential relationship deserves scientific effort for examination so that the importance or place of discourse marker use for learners/non-native speakers can be better demonstrated.

1.3. Aim

In this regard, this study aims to investigate the use of discourse markers by Turkish and native English-speaking British university students in informal interviews, and to determine whether discourse marker use is in fact related to speaking fluency. In line with these aims, the following research questions are addressed in the study:

1. Are there any differences between Turkish and British university students in terms of the frequency and function of discourse markers they employ?
2. Are there any differences between first- and fourth-year Turkish university students in English language teaching (ELT) in terms of the frequency of discourse markers they employ?
3. Is there a relationship between Turkish university students' use of discourse markers and their utterance fluency and perceived fluency?

1.4. Significance of the Study

This study can have significance in terms of providing data regarding the extent to which Turkish EFL learners employ discourse markers in spoken discourse. Such data can be useful to determine what aspects of discourse markers should be addressed and emphasised more in English language instruction in the Turkish context. On the other hand, since the university students to be examined in this study major in English language teaching and will be teachers of English, data on their use of discourse markers can provide insights about how competent Turkish pre-service EFL teachers are in this respect. Moreover, comparing first and final year students' performance can show whether the training they receive during their undergraduate teacher education contributes to their use of discourse markers.

Another aspect of the employment of discourse markers is that they positively contribute to listeners' perception of non-native speakers' fluency through the functions they fulfil such as turn management and holding the floor (Olynyk et al., 1987, p. 123). Although native-like use of discourse markers is argued to be crucial for learners to express themselves in the target language in a fluent way (Hellermann and Vergun, 2007, p. 160; Sankoff et al., 1997, p. 191), whether the use of discourse markers is in

fact a characteristic of fluent language use has not been addressed sufficiently in the literature before. In this regard, this study focuses on the use of discourse markers by Turkish EFL university students in comparison with British university students in informal interviews, and investigate whether frequent and native-like use of discourse markers is related to speaking fluency.

A further possible significance of the present study can be related to the nature of the data collected. In the literature, a considerable amount of the studies on learner language have examined the written mode of English, whereas fewer studies focused on learners' spoken language. This is partly due to the fact that compiling a spoken corpus requires much more time and effort because language samples need to be audio-recorded and then the spoken data have to be manually transcribed. However, it seems to be an area that, considering the limited literature, needs to be further explored. In this sense, this study can also contribute to the literature as it compiles and analyses a spoken corpus of learner language consisting of informal interviews.

1.5. Limitations of the Study

The findings of this study aim to reveal the nature of discourse markers as employed by Turkish and British university students, and of the relationship between discourse marker use and speech fluency. However, the generalisability of these findings can be limited in terms of the genre analysed in this study, the scope of the discourse markers examined, and the characteristics of the participants contributing to both corpora included.

The data analysed in the present study consist of informal interviews. The nature of this genre is more casual and less structured compared to formal interviews. When interpreting and generalising the findings of this study, one should keep in mind that the speech context is limited to informal interviews in which an interviewer who is a research assistant or a lecturer and an interviewee who is a student and may or may not have acquaintance with the interviewer have a conversation around certain topics. Therefore, certain functions of discourse markers that are more frequent in other speech contexts may not appear here. This is not really a significant downside because in such

studies there has to be a restriction regarding the type or context of the dataset since collecting and creating a spoken corpus is a tiresome and time-consuming process.

The discourse markers examined in this study include *so*, *like*, *you know*, *I mean* and *well*. Why or based on what criteria these discourse markers were selected is explained in detail in the method section below. However, the findings should be interpreted by keeping in mind that only these discourse markers were analysed in this study. Considering that conducting a functional analysis of discourse markers requires a considerable amount of time and effort, most studies focus only on certain discourse markers if they have functional analysis in their methodologies.

The characteristics of the participants who were interviewed for both the Turkish corpus and LOCNEC (i.e. British corpus) (De Cock, 2004) are not fully balanced in terms of gender and age. For instance, 32 of the 50 interviewees were female in the Turkish corpus, while 30 of the 50 interviewees were female in the LOCNEC; in other words, most of the interviewees were female. There are different studies that specifically focused on gender differences in the use of discourse markers. For instance, Östman (1981, p. 70) concluded that *you know* tended to be employed more frequently by women than by men although it was not the same in all conversations. As for age, the Turkish interviewees ranged from 18 to 30 (the average was about 21), and the age of the British interviewees was also between 18 and 30 (the average was about 22). For instance, Blyth et al. (1990: 219) report that the "use of *be like* dropped off sharply after the age of 25 and disappeared altogether at the age of 38", which can be regarded as an outdated finding, though. However, it should be noted that the British corpus was bought to be analysed in this study, so the researcher did not have any choice in this respect, while the target population of the Turkish corpus mostly included female students, so it is normal that the sample also includes more female students. To sum up, the limitations mentioned above should be considered while interpreting and generalising the results of this study.

2. LITERATURE REVIEW

In this section, the terminology used to refer to discourse markers in the literature is firstly clarified. Secondly, the definitions of the concept 'discourse markers' are discussed with reference to different approaches in the literature. Thirdly, the characteristics of discourse markers are elaborated, and lastly, empirical studies on the use of discourse markers, particularly by non-native speakers of English, are addressed.

2.1. Terminology

A variety of descriptions were attributed to the phenomenon of discourse markers before large scale studies were conducted. They were described as fumbles, hedges, fillers, evincives, starters, conversation greasers and compromisers (Andersson and Trudgill, 1990). As they obviously have discourse-marking functions, they were also regarded as the main organizers and facilitators of discourse (Levinson, 1983: Schiffrin, 1987: Fuller, 2003). In comprehensive works, the concept of discourse markers has also been termed differently by various researchers. These terms include discourse particles (Aijmer, 2002; Fischer, 2006), connectives (Bazzanella, 1990; Fraser, 1988), pragmatic markers (Erman, 2001; Aijmer, 2005), discourse operators (Redeker, 1991) and discourse markers (Schiffrin, 1987; Blakemore, 2002; Fuller, 2003; Müller, 2005; Fung and Carter, 2007).

When the literature on this concept is thoroughly examined, two of the terms provided above seem to be used more commonly than the others: discourse markers and discourse particles. However, discourse particles were mostly used in studies focusing on languages other than English. This is because 'particle' is used to traditionally refer to a specific grammatical category in certain languages (e.g. modal particles in German and Chinese) (Müller, 2005, p. 3; Mei, 2012, p. 19). This term appeared in studies on, for example, German (Abraham, 1991; Fischer and Drescher, 1996), French (Hasen, 1998) and Chinese (Lee-Wong, 2001). Therefore, the term 'discourse markers' was adopted in the current study since the focus was on spoken English.

2.2. Discourse Markers: History, Definition and Theoretical Approaches

Discourse markers are generally agreed to contribute to the pragmatic meaning of utterances; therefore, they play an important role in speakers' pragmatic competence. However, there seems to be no consensus regarding the definition of discourse markers, or which linguistic items should be considered as discourse markers. More specifically, as Schourup (1999, p. 227) states, fundamental issues that are not agreed upon in the literature include "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". In other words, scholars have viewed discourse markers from different perspectives.

In the literature, one of the early references to discourse markers as a linguistic entity is perhaps that of Labov and Franshel (1977 cited in Fraser, 1999) in which they stated the following with regard to *well*:

> As a discourse marker, *well* refers backwards to some topic that is already shared knowledge among participants. When *well* is the first element in a discourse or a topic, this reference is necessarily to an unstated topic of joint concern. (p. 156)

Later, before comprehensive works were published, Levinson (1983) also provided brief comments on discourse markers in that he considered them as a class that is worth to study, but did not name them. He argued his views 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. Examples are utterance-initial usages of *but, therefore, in conclusion, to the contrary, still, however, anyway, well, besides, actually, all in all, so, after all,* and so on. ... what they seem to do is indicate, often in very complex ways, just how the utterance that contains them is a response to, or a continuation of, some portion of the prior discourse. (pp. 87-88)

Levinson, as is seen, provided sample words and phrases as having certain functions in discourse, but did not elaborate on this beyond these comments. Certain comprehensive and widely cited seminal works were conducted by various scholars in

the following period. These scholars and their works can be argued to represent, at least, three theoretical accounts to the understanding of discourse markers: coherence-based, grammatical-pragmatic, and relevance-based perspectives.

In the coherence-based perspective, the first comprehensive analysis was conducted by Schiffrin (1987), which is also seen as the first and most detailed effort addressing discourse markers. In her early definition, she defines discourse markers as "sequentially dependent elements which bracket units of talk" (p. 31). Here, what she means by 'sequentially dependent' is that the units of talk before and after a discourse marker indicates the social and pragmatic meaning communicated or inferred. In her book, she suggested a general framework of discourse markers, and analysed *and, because, but, I mean, now, oh, or, so, then, well* and *y'know* in unstructured conversations and spontaneous speech. She claims that discourse markers do not easily fit into a single linguistic class; moreover, she also suggests that even paralinguistic features and non-verbal gestures are possible discourse markers.

Schiffrin (1987) mainly focused on the ways in which discourse markers function to "add to discourse coherence" (p. 326). Coherence is "constructed through relations between adjacent units in discourse" (p. 24), and in her framework, she considers discourse markers as contextual coordinates for utterances by locating them on five planes of talk, each with its own type of coherence. These planes of talk include ideational structure, action structure, exchange structure, participation framework and information state. She argues that coherence is established by means of the relations between adjacent units in discourse due to their semantic and syntactic properties. Ideational structure reflects certain relationships between the ideas (propositions) found within the discourse, including cohesive relations, topic relations, and functional relations, whereas action structure reflects the sequence of speech acts which occur within the discourse. Exchange structure reflects the mechanics of the conversational interchange and shows the result of the participant turn-taking and how these alternations are related to each other, while participation framework reflects the ways in which the speakers and hearers can relate to one another as well as orientation toward utterances. Lastly, information state reflects the ongoing organisation and management of knowledge and meta-knowledge as it evolves over the course of the discourse.

According to Schiffrin (1987), discourse markers such as *and, but, or, so* and *because* are operative on the ideational structure, and indicate three types of relations that contribute to the configuration of idea structures: cohesive relations, topic relations and functional relations. Those such as *well, oh, now, y'know* and *I mean* operate on the other levels: exchange, action, participation framework and information state. For example, *oh*, she states, is a marker of information management, and "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" (p. 101). In overall, what she proposes is that discourse markers provide contextual coordinates for an utterance by (i) locating the utterance on one or more planes of talk of her discourse model, (ii) indexing the utterances to speaker, hearer, or both, and (iii) indexing the utterances to prior and/or subsequent discourse. In her account, discourse markers serve an integrative function in discourse and thus contribute to discourse coherence.

Although Schiffrin (1987) published the first comprehensive work on discourse markers, she also attracted some criticism from other scholars. For instance, in her critique of Schiffrin, Redeker (1991) argues that some discrepancies exist between the descriptions of individual markers and the places of these markers in her model (i.e. assigning markers to the planes of talk), and that the planes she proposed in her model are not all comparable, and not well-defined. Regarding the first point, she says that some functions of discourse markers that are not included in Schiffrin (1987) can be found in other works in the literature (p. 1151). As for the second point, she argues that the information structure and participation framework are not the same with other three planes (i.e. ideational structure, action structure and exchange structure) because "the cognitions and attitudes composing those two components concern individual utterances, while the building blocks at the other three planes are relational concepts" (Redeker 1991,p. 1162). As a result, she proposed a broader framework embracing all connective expressions. She labelled them as 'discourse operators' instead of discourse markers and provided the following definition:

> A *discourse operator* is any expression –a word or phrase- that is used with the primary function of bringing to the listener's attention to a particular kind of relation between the discourse unit it introduces and the immediate discourse context. An utterance in this definition is an intonationally and structurally bounded, usually clausal unit. (p. 1168)

Redeker presented examples of what are not discourse markers: clausal indicators of discourse structure (e.g. *let me tell you a story, as I said before, since this is so)*, deictic expressions as far as they are not used anaphorically (e.g. *now, here, today)*, anaphoric pronouns and noun phrases, and any expressions whose scope does not exhaust the utterance (1991, p. 1168). On the other hand, she also concluded that among Schiffrin's five planes of talk, information structure and participation framework are not independent of the other three planes, and therefore, should be incorporated into them. She thus provided a revised model of discourse coherence based on three components: ideational structure and rhetorical structure, which are somewhat equivalent to Schiffrin's ideational structure and action structure, and a sequential structure, which is somewhat equivalent to an extended version of Schiffrin's exchange structure. She points out that any utterance in discourse is considered to always participate in all three components, however, one of these components will be usually dominant. Redeker's definition was also shared in that of discourse connective in Blakemore (1987) which will be mentioned below.

As for the second theoretical account, Fraser (1990, 1996, 1999) took a grammatical-pragmatic perspective of discourse markers, which he labelled 'pragmatic markers'. His definition was slightly different from that of Schiffrin, and included vocalisations such as *oh* and *ah*. According to him, discourse markers are linguistic expressions signalling an intended relationship between the utterance introduced by a discourse marker, and the preceding utterance. He characterised a discourse marker as a linguistic expression only (unlike Schiffrin who argues that paralinguistic features and non-verbal gestures are possible discourse markers) which: (i) has a core meaning which can be enriched by the context, and (ii) signals the relationship that the speaker intends between the utterance the discourse marker introduces and the foregoing utterance (rather than only highlighting the relationship, as Schiffrin suggests) (Fraser, 1999).

Fraser (1999) defines discourse markers as a class of lexical expressions drawn primarily from the syntactic classes of conjunctions, adverbs, and prepositional phrases. Compared to Schiffrin's coherence model, Fraser seems to contribute to a more complete generalisation and a pragmatic view towards different markers within his grammatical-pragmatic perspective. Thus, he suggests that discourse markers do not only ensure textual coherence as in what the coherence model argues, but also signal the speakers' intention to the next turn in the preceding utterance.

The third theoretical perspective is the relevance theory framework, which is mainly advocated by Blakemore (1987, 1992) who was studying discourse markers almost at the same time with Schiffrin (1987). Blakemore focused on how discourse markers, which he actually preferred to call 'discourse connectives', impose constraints on implicatures. More specifically he defined them as "expressions that constrain the interpretation of the utterances that contain them by virtue of the inferential connections they express" (Blakemore 1987, p. 105). According to Blakemore, one draws inferences based on the relevance of his/her assumptions and the contexts in understanding an utterance, and this process is constrained by discourse markers. She provides the following example:

> [On seeing someone carrying lots of parcels]
> So you've spent all your money. (Blakemore, 1988, p. 188)

In this example, the utterance *so you've spent all your money* is produced based on what the speaker has seen, while the hearer's comprehension is constrained by the discourse marker *so*; accordingly, the speaker is drawing a conclusion after he has seen the listener's parcels. Apparently, *so* does not have a linguistic antecedent in this example, but interprets the following utterance in a way that is relevant to the situational context.

Blakemore then proposes that discourse connectives do not have a representational meaning the way lexical expressions have, but only have a procedural meaning, which consists of instructions about how to manipulate the conceptual representation of the utterance (Blakemore, 1992). In other words, he claims that

discourse connectives merely have a procedural meaning and are limited to a specific context. The meaning of discourse connectives can be understood, as he argues, by looking at the context presentation of when they are uttered in discourse. Similar to Fraser (1999), he suggests that discourse connectives do not contribute to a representational meaning, but rather to a procedural and pragmatic meaning, which encodes instructions for processing propositional representations of the utterances (Blakemore, 1992). Within this procedural meaning, unlike Schriffin (1987), both Fraser (1999) and Blakemore (1992) agree that discourse markers mainly focus on the way communication is negotiated between the speaker and listener in the discourse, rather than on its content.

Another scholar who also adopts a relevance theory approach, Jucker (1993) believes that it 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). He provides the following three aspects of the theory that should be considered to understand how the notion of context is used in it:

> 1. Every utterance comes with a guaranteee 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, p. 438)

Therefore, discourse coherence is established when an utterance is recognised as relevant by the hearer in the context provided by the immediately preceding/following utterance. Additionally, Jucker (1993) argues that cohesive devices can ensure coherence since it is a function of utterance interpretation, and thus "relevance theory provides more plausible explanations for a wide range of occurences of discourse markers and it is superior as it accounts for all the examples in the relevant literature" (p.440). Above all, discourse connectives, as in Blakemore's terms, do not contribute to the proposition expressed by an utterance, but point the hearer to the context in which he/she is expected to process the utterance as well as to the conclusions he/she should be drawing from it (Rouchota, 1996: 5).

Other than the three theoretical accounts of discourse markers discussed above, more recent works provided definitions and explanations based on these accounts as well as the corpus-based evidence they reported. In a more recent work, Aijmer (2002) defined discourse markers as "a class of words with unique formal, functional and pragmatic properties" (p. 2). She took into account that indexical potential of discourse markers and argued that they are indexed to attitudes, to participants and to the text. Consequently, they have discourse functions both on the textual and interpersonal level, and need to be described in terms of discourse contexts that extend beyond turn boundaries (Aijmer, 2002). On the other hand, based on both Schiffrin's (1987) model of coherence and Aijmer's (2002) interpersonal perspective, Fung and Carter (2007) argued that an exclusive emphasis on textual coherence and relevance is not sufficient, and both textual and interpersonal dimensions of discourse markers should be considered. Thus, their theoretical framework adopted a functionally-based account, and categorised discourse markers under functional headings that include interpersonal (i.e. marks the affective and social functions), referential (i.e. marks relationships between verbal activities preceding and following a discourse marker), structural (i.e. indicates the discourse in progress) and cognitive (i.e. provides information about the cognitive state of speakers, e.g. reformulating, elaborating, or marking hesitation) categories.

Apart from the accounts described above, some scholars (e.g. Nikula, 1996; Östman, 1981) approached the notion of discourse markers from the perspective of the politeness theory. For instance, Östman (1981) examined *you know* in this framework. He argues that the speaker uses *you know* because he/she wants the hearer not to argue with him/her, but to cooperate and accept the content of the speaker's utterance. Because the hearer's opposition to the content of that utterance can be a threat to the speaker's face, *you know* here operates as a face-saving device (pp. 17-22). Östman (1981, p.19) also provides other ways in which *you know* can contribute to politeness. For instance, by using *you know*, the speaker pretends shared knowledge, which achieves intimacy. He exemplifies this explanation: after an hour of teaching, a professor starts to frequently use *you know* to show that he switches to a more informal style and ensures a more equal or balanced, and after ten minutes, the students accept this style and do the same relationship. There is another way in which you know fulfils

politeness functions. By using it, the speaker pretends shared knowledge, and this achieves intimacy.

2.3. Properties of Discourse Markers

With respect to the properties of discourse markers, Schiffrin (1987) states that discourse markers are independent of sentential structure, and thus, it is difficult to syntactically define their locations within a sentence. Here, she refers to two properties: syntactic independence and place of occurrence in relation to sentence structure. To clarify these properties, a discourse marker is not bound to the sentence structure, yet stays "outside the syntactic structure" (Erman, 2001, p. 1339), and can occur in sentence initial, medial and final positions (Schiffrin, 1987; Brinton, 1996). Moreover, discourse markers lack semantic content, which means that they have "little or no meaning" (Erman, 2001, p. 1339) and "relatively little semantic content" (Simon-Vanderbergen, 2001, p. 82). Multifunctionality is another property of discourse markers and refers to serving more than one pragmatic function in different instances or in a particular instance (Lam, 2008). With respect to their functions, discourse markers can fulfil a variety of functions in discourse. Brinton (1990) provides a list of general functions that discourse markers fulfil:

- initiate discourse
- mark a boundary in discourse, i.e. to indicate a shift or partial shift in topic
- preface a response or a reaction
- serve as a filler or delaying tactic
- aid the speaker in holding the floor
- effect an interaction or sharing between speaker and hearer
- bracket the discourse either cataphorically or anaphorically
- mark either foregrounded or backgrounded information (Brinton, 1990, pp. 47-48)

Following the broad categories in Halliday and Hassan (1976, pp. 26-28), in her later work, Brinton (1998) proposed that the functions presented above can be classified

into two broad groups: textual and interpersonal. This grouping was also adopted in Aijmer (2002) and later in Buysse (2010) who both argued that discourse markers function in one of two levels, being structural or textual, and interpersonal or phatic. While textual function plays a role in the organization of discourse, interpersonal functions are more interactional. Textual functions include getting the hearer's attention, initiating and ending discourse, sustaining discourse, marking boundaries (topic shifts and episode boundaries), constraining the relevance of adjoining clauses and repairing discourse. On the other hand, the interpersonal category has subjective functions such as expressing response, reactions, attitudes, understanding, tentativeness, or continued attention, and interactive functions such as expressing intimacy, cooperation, shared knowledge, deference, or face-saving (politeness) (Brinton, 1998, p. 12). However, it is not possible to make a clear cut distinction regarding which discourse marker can be regarded in which of these categories. All discourse markers perform a variety of functions. The same marker can contribute to discourse in one of these categories in an instance, and in the other category in another instance (Buysse, 2010).

In addition to the two categories in Brinton's framework, i.e. textual and interpersonal, most studies further include a third domain: interactional category (e.g. Müller, 2005; Lam, 2010). Interactional functions are mostly associated with the speaker's planning process and turn-managing activities. They create an area of sharing and intimacy by helping the speaker to hint to the hearer to interact and in this way create common knowledge (Müller, 2005, p.68). Thus, they ensure the smooth flow of communications through such functions. Moreover, discourse markers with interactional functions can help speakers overcome difficulties in their planning process. For instance, when speakers search for the right word or structure, they can employ expressions such as *I mean* and *you know* for time stalling purposes so that their processing phase can be lengthened, or the hearer may contribute to the interaction and help them out (Erman, 2001). Interactional functions also facilitate interactions through turn-taking or floor-holding (Aijmer, 2002). Such elements can be used when the hearer intends to take the floor from the speaker or to hold the floor to speak when interrupted.

2.4. Discourse Markers and Second Language Acquisition

Given the definition and characteristics of discourse markers, work has also been done on how these lexical items are acquired by learners in the contexts of different languages such as English, French, Spanish and Japanese. In an earlier study, Hays (1992) examined the employment of different types of discourse markers by Japanese learners of English. He reported that certain markers such as 'but', 'and', and 'so' were used frequently, 'well' and 'you know' were used by very few learners. He then argued that there is a developmental order for learners' acquisition of discourse markers, and those on the ideational plane are taught and used first, but those that are more pragmatic appear later in learner speech. In a later work, Andersen et al. (1999) stated that textual functions of discourse markers are acquired somewhat earlier than their register functions. Based on both American and French role-play data, they also concluded that the social meanings of these forms develop at a later stage. Similarly, Romero Trillo (2002) stated that learners follow a formal and pragmatic track in their linguistic development, in that the formal track relates to the grammatical and semantic rules while the pragmatic track relates to the social use of language in various contexts and registers. He further argued that native speakers develop both tracks simultaneously through contact with natural language, but non-native learners in a non-target language environment develop these two tracks by means of formal instruction. Therefore, the pragmatic track is more difficult since it is linked to cognitive, affective and socio-cultural meanings expressed by language forms, which is also difficult to integrate in educational curricula. As a result, he also argued that non-native speakers find the formal-functional dichotomy problematic in their learning process due to the nature of their learning environment being not contextualised, which lead to an acquisition of a simplified and context-free register of the target language with, in most cases, no explicit relationship between form and function.

Research also focused on how discourse markers should be presented to learners in a classroom context. Hernandez (2008) examined the effect of explicit instruction and input flood on students' use of Spanish discourse markers on a simulated oral proficiency interview. Explicit instruction combined with input flood was found to be more effective than input flood alone in promoting students' use of discourse markers.

Later, in a further study, Hernandez (2011) investigated the combined effect of explicit instruction and input flood vs. input flood alone on students' use of discourse markers to narrate a past event. According to the results, the combined effect of explicit instruction and input flood was not superior to input flood alone in promoting learners' use of discourse markers. In a more recent study, Kim (2016) explored the effect of a pragmatically-oriented intervention, which introduced a unified concept of *ndesu* that is a Japanese discourse marker, in which the pragmatic aspect was emphasised in comparison to a textbook-based approach based on several representative functions of this discourse marker with no mention of how each function is related to each other. The results yielded a positive, immediate effect for learning *ndesu*; in other words, explicit *ndesu* instruction was effective regardless of the type of instruction (i.e. pragmatically-oriented or textbook-based). Yet, the pragmatically-oriented group performed significantly better than the textbook-based group in the posttest. In other words, an approach emphasising the pragmatic aspect of this discourse marker was more effective than merely presenting certain representative functions through a textbook.

To sum up, research that focuses on learners' acquisition of discourse markers shows that non-textual discourse markers, or those that are more pragmatic, are acquired later than textual markers or functions. As for classroom instruction, explicit instruction of discourse markers with an emphasis on their pragmatic or non-textual functions can be more effective compared to merely adopting an input-enhanced approach.

2.5. Empirical Studies of the Use of Discourse Markers

Having mentioned the terminology, properties and functions regarding the concept of discourse markers, it can be asserted that it has been a topic of interest by researchers in pragmatics for almost three decades. It has been studied from a wide range of aspects including descriptive functional approaches (Schiffrin, 1987; Fraser, 1990), taxonomic discussions (Andersen, 2001) and gender differences in the use of discourse markers (Erman, 1992). All of these studies agree that discourse markers contribute to pragmatic meaning in language use, and play a significant role in speakers' pragmatic competence. Discourse markers are like "the oil which helps us perform the complex task of

spontaneous speech production and interaction smoothly and efficiently" (Crystal, 1988, p. 48).

2.5.1. Empirical studies on non-native speakers of English

Most of the early studies on discourse markers focused on the language produced by native speakers, and studies on non-native speaker discourse have existed in the literature only for the last decade. The use of discourse markers by non-native speakers has been studied from different aspects in various contexts. Table 2.1 presents a brief summary of some of the recent studies on the use of discourse markers by non-native speakers.

Table 2.1. *Recent studies on the use of discourse markers (DMs) by non-native speakers*

Study	Context	Focus
Romero Trillo (2002)	Spanish EFL	Pragmatic fossilization of DMs
Müller (2005)	German EFL	Use of four DMs in various spoken corpora
Fung and Carter (2007)	Chinese EFL	Use of DMs in classroom interactions
Hellerman and Vergun (2007)	Adult ESL Learners	Use of three DMs in classroom interactions and at home interviews
Liao (2009)	Taiwanese and Chinese ESL	Use of nine discourse markers lecture discussions and interviews
Buysse (2010)	Flemish (EFL)	Use of six discourse markers in informal interviews
Polat (2011)	Turkish (ESL)	Use of three DMs in informal interviews by a Turkish immigrant
Bu (2013)	Chinese (EFL)	Use of nine DMs in classroom discussions and interviews
Mei (2012)	Chinese (EFL)	Use of two DMs in various spoken corpora
Aşık and Cephe (2013)	Turkish (EFL)	Use of DMs in in-class presentations
Buysse (2015)	Dutch, French, German, Spanish, Chinese, Swedish EFL	Use of *well* in informal interviews
Buysse (2017)	Dutch, French, German, Spanish EFL	Use of *you know* in informal interviews
Diskin (2017)	Polish & Chinese (ESL)	Use of *like* in sociolinguistic interviews

Zorluel Özer and Okan (2018)	Turkish (EFL)	Use of DMs in teacher lectures in EFL classrooms

Some of these studies compared the use of discourse markers by non-native speakers and native speakers of English (Romero Trillo, 2002; Müller, 2005; Fung and Carter, 2007; Buysse, 2010, 2015, 2017; Mei, 2012; Aşık and Cephe, 2013; Diskin, 2017; Zoruluer Özer and Okan, 2018) while others focused merely on non-native speaker performance (Hellerman and Vergun, 2007; Liao, 2009; Polat, 2011; Bu, 2013). Romero Trillo (2002), one of the early studies in the literature on non-native speaker discourse, investigated the use of six discourse markers *look*, *listen*, *you know*, *I mean*, *well* and *you see* by native speaker and Spanish EFL learner children and adults. The results revealed that native speakers increase their use of involvement markers, which are mostly concerned with the argumentative process and the social and cognitive relationship between speaker and hearer, as they grow up while non-native speakers do not. However, non-native speakers are able to use operative markers, which are mostly related to the mechanics of interactions, in a way similar to native speakers. In a later study, Müller (2005) examined the use of *well, you know, like* and *so* in German EFL learners' conversation in comparison to that of native speakers of English. The results of his analysis showed that although German speakers employed the four discourse markers, there were differences in the use of the individual functions. Some functions used by native speakers were almost completely unknown to German speakers while some functions were only used by Germans.

Focusing on classroom interactions, Fung and Carter (2007) compared the use of discourse markers by learners in Hong Kong and British native speakers. The Hong Kong learners demonstrated an abundant use of referentially functional discourse markers (e.g. *and, but, because, OK* and *so*), but a limited use of other markers including *yeah, really, say, sort of, I see, you see* and *well* while native speakers used a wider range of pragmatic functions. Hellerman and Vergun (2007) focused on adult learners' use of discourse markers in classroom interactions and at home interviews. The participants who did not have any previous formal second language instruction rarely used three discourse markers, i.e. *well*, *you know*, and *like*, and those who were

more acculturated to the United States exhibited higher use of these discourse markers. Liao (2009) studied Chinese and Taiwanese learners in an ESL context in the US. Her participants were teaching assistants at a university and the data consisted of their lecture discussions and interviews conducted with them. In general, the participants' use of nine focal discourse markers was comparable to native speakers (i.e. as reported in previous studies), but they performed an infrequent use of *well* and *I mean* in overall. In another study on Chinese learners, Mei (2012) investigated how *I mean* and *you know* are used by Chinese EFL learners as compared to British speakers. *I mean* was markedly underused by the Chinese learners, but they used most of its functions appeared in the native speaker data. This difference was argued to stem from the learners' L1 and the Chinese equivalent of *I mean* was not infrequent in language use and did not have a pragmatic function. As for *you know*, it was markedly overused by the Chinese learners compared to the native speakers, and this difference was also explained by L1 influence.

Some of the studies in the literature (Buysse, 2010, 2015, 2017; Polat, 2011) examined non-native speakers' use of discourse markers in informal interviews. Buysse (2010) focused on six discourse markers, namely *so, well, you know, like, kind of/kinda/ sort of* and *I mean* in interviews with Flemish and British university students in linguistics and language teaching. While Flemish students seem to have slightly overused *so* and *well*, they significantly underused *you know, like, kind of/kinda/sort of* and *I mean* compared to British students. He explains this difference by referring to the categories of textual and interpersonal functions proposed by Brinton (1998) and Aijmer (2002). This means that *so* and *well* were more typically textual markers compared to *you know, like, kind of/kinda/sort of* and *I mean*, which were more typically interpersonal markers. Buysse then conducted two further studies on two specific discourse markers *well* (2015) and *you know* (2017). In Buysse (2015), he examined the use of *well* in the Dutch, French, German, Spanish and Chinese components of the Louvain International Database of Spoken English Interlanguage (LINDSEI) and compared these results with Aijmer's (2011) findings for the Swedish component of LINDSEI and a comparable native speaker corpus (i.e. LOCNEC). All learner groups used *well* significantly more frequently than their native peers; the Chinese learners who

used it extremely less frequently were an exception. The learners employed *well* in functions that are mostly related to speech management. On the other hand, in Buysse (2017), the researcher this time focused on *you know* in the Dutch, French, German and Spanish components of the LINDSEI and its native speaker equivalent, LOCNEC. All learner groups used *you know* considerably less frequently compared to the native speakers. In these two studies (2015, 2017), Buysse also conducted a functional analysis, and compared and reported the findings in the two corpora in this respect. In another study based on interviews, Polat (2011) interviewed a Turkish immigrant in the US over a one-year period and examined the use of *you know, like* and *well*. The participant used *you know* extremely frequently during the interviews while he employed *like* to varying degrees throughout the data gathering process, but never used *well* as a discourse marker. The participant's explanation for overusing *you know* was that he wanted to make sure the interviewer understood him because he did not trust his English. To sum up the results of the previous studies, non-native speakers show a clear difference in their use of discourse markers from native speakers, and this difference varies depending on the L1 and the focus of studies.

Although discourse marker use has been studied in various contexts, the literature on the Turkish context has been extremely rare. There are two studies that examined the use of discourse markers in the speech of Turkish speakers of English (Aşık and Cephe, 2013; Zorluel Özer and Okan, 2018). Aşık and Cephe (2013) investigated the use of discourse markers in the conversation of Turkish EFL learners in comparison with a native speaker corpus. The learners were pre-service EFL teachers studying at a university and their presentations for two courses, Sociolinguistics and Language Teaching, and Pragmatics and Language Teaching, comprised the Turkish learner corpus in the study. The data were manually examined to determine discourse markers used by the Turkish participants, and their occurrences in their spoken English discourse by comparing them with the ones used in native speakers' spoken discourse and their use of functions. With respect to raw occurrences, the native speakers' use was more frequent, and besides, they used much more different markers with several functions. The researchers explained this difference by arguing that the Turkish learners had not been exposed to such items while learning English as a foreign language through

instructional materials or their language teachers. Their study merely focused on planned or scripted language produced by the students and investigated a number of words and expressions that were identified as discourse markers. Another study by Zorluel Özer and Okan (2018) examined the discourse markers used by Turkish teachers and native teachers in EFL classrooms, and comparing them in terms of variety and frequency. The data they analysed consisted of audio-recordings collected from two Turkish teachers working at a Turkish university and two native English speaking EFL teachers' lectures working at a language school in London, UK. They were mainly interested in teachers' use of discourse markers because they argued that using such items could improve students' comprehension of lectures. The results showed that the Turkish teachers used 29 different discourse markers, while the native teachers used 37 different discourse markers; so, the Turkish teachers employed discourse markers with a less variety. It was also found that the Turkish teachers significantly underused most discourse markers when compared to their native counterparts. Based on these results, the researchers claim that it is necessary to raise non-native English language teachers' awareness towards the significance of discourse markers in the spoken discourse of native speakers. However, their study included data collected from only four individuals who teach English in a classroom setting, which may provide insights about teachers' competence in this issue, but the language that a teacher uses in such context can be monologic at most times, and their performance may differ in a more dialogic atmosphere. As a result, it can be argued that there is still a need for studies that would examine Turkish EFL learners' use of discourse markers in more natural or spontaneous speech to reveal to what extent they approximate to native speaker performance.

2.5.2. Discourse marker use and fluency

Another key construct to be focused in this study is speaking fluency that is briefly described as the speed and smoothness of oral performance. A detailed working definition would be the rapid, smooth, accurate, lucid, and efficient translation of thought or communicative intention into language under the temporal constraints of on-line processing (Lennon, 2000, p.26). However, Segalowitz (2010) makes a distinction between three notions of fluency that include cognitive fluency, utterance fluency and

perceived fluency. Cognitive fluency is about speakers' abilities to efficiently plan and execute their speech. Utterance fluency refers to the fluency that is measured by means of the temporal aspects of a speech sample, and it has three other aspects within itself: breakdown fluency, speed fluency, and repair fluency (Skehan, 2003; Tavakoli and Skehan, 2005). Breakdown fluency is related to the ongoing flow of speech, which can be measured by counting the number and length of filled and unfilled pauses. Speed fluency that is measured by calculating speech rate (e.g. number of syllables per second) is about the speed at which speech is delivered. Repair fluency is the frequency of false starts speakers use, corrections they make and repetitions they produce. As for the third notion of fluency, i.e. perceived fluency, it is related to the impression that listeners have of the fluency of a speech sample (De Jong, Steinel, Florijn, Schoonen and Hulstijn, 2013). In the literature, a number of temporal variables such as speech rate, articulation rate, mean length of runs, and number and mean length of pauses per minute, which contribute to utterance fluency, are used to evaluate overall fluency of oral performance along with the perceived ratings of native speakers, which is evaluated as perceived fluency (Kormos and Denes, 2004).

Since discourse markers are argued to play a significant role in the pragmatic competence and spoken interaction of a speaker is discourse markers (Müller, 2005; Mullan, 2010), speakers can be perceived as dysfluent when they perform non-target-like use of discourse markers (Hellermann and Vergun, 2007). Consequently, discourse marker use, which native speakers perform effortlessly, is also necessary for learners to express themselves in a fluent and confident way in that language (Sankoff et al., 1997). The employment of discourse markers can indicate more advanced and successful use of language because when individuals use discourse markers, they have greater involvement in the ongoing communication, attend to situational needs, manage the social and cognitive relationship with the addressee, get the floor, avoid gaps within interaction, fill the pose they need for cognitive processing, and have more skilful turn management (Wei, 2011). These advantages of individuals' use of discourse markers may indeed bring a higher level of fluency in their speech.

With respect to the relationship between the use of discourse markers and individuals' oral fluency, Olynyk, d'Anglejan and Sankoff (1987) argued that the use of

discourse markers, which they called speech markers, can be seen as a positive contributor to listeners' perception of non-native speakers' fluency. In a study they conducted with French speakers' L2 English in a military context, they compared individuals of high fluency with those of low fluency. Their study reported that more discourse markers were employed by the high fluency group. Similar findings were also obtained by Sankoff et al. (1997) who examined how speakers of English use discourse markers in their L2 French. Their results showed that "very low-frequency use of discourse markers is a signal of lesser overall linguistic competence, whereas higher frequency is the hallmark of the fluent speaker" (Sankoff et al., 1997: 204). They thus suggested that "the ability to express oneself fluently and confidently in a second language entails the use of those discourse markers that native speakers produce so effortlessly" (Sankoff et al., 1997: 214).

Although it is self-evident that the use of discourse markers can contribute to non-native English speakers' oral fluency (Mei, 2012), the relationship between the use of these markers and fluency has not attracted much attention as a direct topic of examination in the literature. Exceptions are Olynyk et al. (1987) and Götz (2013). For Olynyk et al. (1987), there are two aspects that can be evaluated as downsides of their study. Firstly, they focused on learners in a very specific context, i.e. a military college in Canada where individuals are nearly English-French bilinguals. Secondly, their indicator of fluency was merely Olynyk's personal observations of the participants over a period of several months in test situations, classrooms and informal contexts, which seems a little subjective and lacks temporal variables that have been widely used in the literature in more recent studies (Kormos and Denes, 2004). On the other hand, Götz (2013) conducted a rather comprehensive study on temporal and perceptive fluency of native and non-native speakers, and the relationship of these two dimensions of fluency with speech management strategies (e.g. pauses, repeats) and other features including discourse markers. She reported that the use of a high proportion of discourse markers did not have a positive impact on the temporal fluency performance of native and non-native speakers. She also looked at perceived fluency based on native speaker ratings, but she did not specifically include the use of discourse markers in that analysis. Consequently, there is a need for further studies that examine the relationship between

non-native English speakers' use of discourse markers and their oral fluency by taking into account both utterance fluency with temporal indicators such as speech rate and articulation rate, and perceived fluency ratings of native speakers of English.

3. METHOD

3.1. Selection of Discourse Markers

In the literature, a wide range of discourse markers have been studied in various studies so far. These studies sought after a theoretical/functional description of discourse markers in general, or a comparison of their use in different registers, task types and groups of speakers. There have been 31 discourse markers as a topic of examination in studies on spoken English discourse. A list of these discourse markers is presented in Table 3.1 below.

Table 3.1. *A list of discourse markers that appeared in the literature*

Discourse Marker	Studies (e.g.)	Appearance
well	Östman (1982), Schourup (1985), Schiffrin (1987), Fraser (1990), Redeker (1990), Romero Trillo (2002), Taguchi (2002), Müller (2005), Fung and Carter (2007), Hellerman and Vergun (2007), Liao (2009), Buysse (2010), Polat (2011), Gilquin (2016)	14
you know	Östman (1982), Schourup (1985), Schiffrin (1987), Redeker (1990), Romero Trillo (2002), Taguchi (2002), Müller (2005), Fung and Carter (2007), Hellerman and Vergun (2007), Liao (2009), Buysse (2010), Polat (2011), Mei (2012), Gilquin (2016)	14
I mean	Östman (1982), Schourup (1985), Schiffrin (1987), Redeker (1990), Romero Trillo (2002), Taguchi (2002), Liao (2009), Buysse (2010), Mei (2012), Gilquin (2016)	10
like	Östman (1982), Schourup (1985), Müller (2005), Fung and Carter (2007), Hellerman and Vergun (2007), Liao (2009), Buysse (2010), Polat (2011), Gilquin (2016)	9
oh	Östman (1982), Schourup (1985), Schiffrin (1987), Redeker (1990), Taguchi (2002), Fung and Carter (2007), Liao (2009)	7
so	Schiffrin (1987), Fraser (1990), Redeker (1990), Müller (2005), Fung and Carter (2007), Buysse (2010), Gilquin (2016)	7
now	Östman (1982), Schourup (1985), Schiffrin (1987), Fraser (1990), Taguchi (2002), Fung and Carter (2007)	6
kind of/kinda/ sort of	Östman (1982), Schourup (1985), Taguchi (2002), Fung and Carter (2007), Buysse (2010), Gilquin (2016)	6
but	Östman (1982), Schiffrin (1987), Fraser (1990), Redeker (1990), Fung and Carter (2007)	5
okay	Fraser (1990), Redeker (1990), Taguchi (2002), Fung and Carter (2007), Liao (2009)	5
and	Östman (1982), Schiffrin (1987), Fung and Carter (2007)	3
or	Schiffrin (1987), Fraser (1990), Fung and Carter (2007)	3
because/cos	Schiffrin (1987), Redeker (1990), Fung and Carter (2007)	3

then	Schiffrin (1987), Fraser (1990), Gilquin (2016)	3
you see	Fraser (1990), Romero Trillo (2002), Taguchi (2002)	3
yeah	Taguchi (2002), Fung and Carter (2007), Liao (2009)	3
right	Taguchi (2002), Fung and Carter (2007), Liao (2009)	3
ah	Schourup (1985), Fraser (1990), Redeker (1990)	3
look	Romero Trillo (2002), Taguchi (2002)	2
actually	Fung and Carter (2007), Liao (2009)	2
really	Taguchi (2002), Fung and Carter (2007)	2
uh	Östman (1982), Schourup (1985)	2
mind you	Schourup (1985), Redeker (1990)	2
all right	Fraser (1990), Redeker (1990)	2
listen	Romero Trillo (2002)	1
just	Fung and Carter (2007)	1
I think	Fung and Carter (2007)	1
yes	Fung and Carter (2007)	1
see	Fung and Carter (2007)	1
say	Fung and Carter (2007)	1
anyway	Fraser (1990)	1

As is seen in Table 3.1, *well* and *you know* were the discourse markers that attracted the most attention in the literature. These two markers were followed by *I mean*, *like*, *oh* and *so*. Due to slight differences in their definitions, some researchers regarded some of the words/phrases in this list as not having discourse marking functions, while other researchers examined them as discourse markers. The reason why *well, you know*, *I mean*, *like*, *oh* and *so* attracted more attention in the literature could be that they have wider range of functions and occur with high frequencies in spoken discourse.

The studies cited in Table 3.1 include both those that focus on a theoretical/ functional description of discourse markers, and empirical efforts that compare the use of discourse markers, for example, by native speakers and non-native speakers of English. If we were to have a closer look on which discourse markers were focused in the latter, it would be useful to compile a list of the lexical items examined as discourse markers in these empirical studies, which is presented in Table 3.2 below.

Table 3.2. *A list of discourse markers examined in studies on native and non-native speakers*

DMs	Study	Data	Selection of DMs
I mean, listen, look, well, you know, you see	Romero Trillo (2002)	naturally occurring conversations of children and adults in L1 & L2 English Ss	different uses in different languages; high frequency in NS corpus
like, so , well, you know	Müller (2005)	narratives and discussions by L1 & L2 English Ss based on a silent movie	high frequency in corpus for meaningful quantitative analysis; a range of diff. characteristics > relevant for use by EFL Ss
actually, and, because, but, cos, I mean, I think, just, like, now, oh, okay, or, really, right, say, see, so, sort of, well, yeah, yes, you know	Fung and Carter (2007)	group discussions for Hong Kong Ss; pedagogic sub-corpus of CANCODE (NS); all pedagogical settings	top 100 most frequent words/ phrases retrieved; 23 items with roles identified with those of DMs were selected
like, well, you know	Hellerman and Vergun (2007)	classroom interactions in an ESL setting	discussed in previous research; high frequency in the dataset
actually, I mean, like, oh, ok, right, well, yeah, you know	Liao (2009)	Chinese TA-led discussions; socio-linguistic interviews	high frequency in a NS corpus
I mean, kind of/ kinda/sort of, like, so, well, you know	Buysse (2010)	informal interviews in L1 & L2 English	most frequently studied in the literature
like, well, you know	Polat (2011)	informal conversations with a NNS in English	most frequently studied in the literature
I mean, you know	Mei (2012)	spontaneous dialogues by NSs; recordings of spoken texts by Chinese Ss	most frequent in spoken English; exclusively discussed in the literature; also frequent in CH learner data; difficult to identify their functions
all words&phrases qualified as DMs	Aşık and Cephe (2013)	TR student presentations in two undergraduate classes; student presentations by American students	functions reported in the literature used to retrieve words/phrases with DM functions
I mean, like, so, sort of, then, well, you know	Gilquin (2016)	informal interviews in L1 & L2 English	commonly cited in the literature

As is seen in Table 3.2, discourse markers are usually selected for examination due to their high frequency in discourse (e.g. Romero Trillo, 2002; Müller, 2005; Liao, 2009) and/or because they are extensively studied in the literature making it easy to compare findings (e.g. Polat, 2011; Mei, 2012; Gilquin, 2016). However, some studies

adopted a purely data-driven approach by either compiling a list of all discourse markers within a reference corpus (as in Fung and Carter, 2007), or deriving all lexical items that function as discourse markers in a research corpus (as in Aşık and Cephe, 2013). In this regard, this study focuses on five discourse markers, namely *so, well, you know, I mean* and *like,* that are commonly studied in the literature (Mei, 2012; Hellerman and Vergun, 2007; Müller, 2005; Romero Trillo, 2002) and also found to be frequent in informal interviews with native English speakers (Buysse, 2010).

3.1.1. *So*

So basically has a resultative or inferential meaning, but it also has various other functions in discourse. In Schiffrin (1987), it is reported to function as marking a result or consequence, taking the discourse back to the main topic after a digression, marking a discourse boundary, and marking a potential turn-transition. Müller (2005) provides a comprehensive picture of *so* functioning in the textual level and interactional level. In the textual level, it functions as marking result or consequence, main idea unit marker, summarising / rewording / giving an example, sequential *so*, and boundary marker, whereas in the interactional level as question or request, opinion, marking implied result, and marker of a transition relevance place.

3.1.2. *Well*

Well has two main functions that are marking an answer to a question, and self-correction or rephrasing (Buysse, 2010). Besides, it can also function as signalling the beginning of a quotation, indicating the change of topic, stalling for time, contradicting, evaluating, and indirectly replying to previous turns (Polat, 2011; Müller, 2005).

3.1.3. *You know*

You know is a discourse marker that is versatile and difficult to describe as it is reported to have up to 30 functions in the literature (Polat, 2011). Among the prominent functions of *you know* are assumption-correction, introducing background information,

reformulation, exemplification, seeking confirmation, summarisation, hesitation marker, and introducing a new topic (Mei, 2012).

3.1.4. *I mean*

According to Schiffrin (1987), *I mean* has three broad functions that are modifier of the speaker's ideas, modifier of the speaker's attentions, and replacement repair. More specific functions also include reformulation, exemplification, cause, reason, result, summarisation, indicating speaker attitude, assumption-correction, hesitation marker, and restart (Brinton, 2007; Mei, 2012).

3.1.5. *Like*

Like is a discourse marker that functions only in the textual level marking new information or focus and indicating approximation (Dailey-O'Cain, 2000; Fuller, 2003). More specifically, it functions as searching for the appropriate expression, marking an approximate number or quantity, introducing an example, and marking lexical focus (Müller, 2005).

3.2. Research Corpora

In this study, two corpora were analysed in accordance with the research questions. The first corpus consists of informal interviews with Turkish speakers of English, which was specifically complied for the purposes of the present study, while the second corpus is the Louvain Corpus of Native English Conversations (LOCNEC) (De Cock, 2004), which was compiled in the scope of a project at the Centre for English Corpus Linguistics, Université Catholique de Louvain.

3.2.1. The Turkish corpus of informal interviews

This corpus was compiled at Anadolu University by means of informal interviews with fifty undergraduate students studying English language teaching at Anadolu Universiy. The students were selected among first year (n=25) and fourth year (n=25) students in the department based on purposive sampling. Starting and final year students

were selected with the aim to see how aware or able pre-service EFL teachers come to a teacher training program and they graduate from it with respect to their use of discourse markers in spoken discourse. Three criteria were considered in the selection of the participants: speaking course grade, GPA and study-abroad experience. For first year students, the participants were selected among those having a course grade of at least BA (i.e. third highest grade between AA and FF in the institution) in the Oral Communication Skills course that they take in the first semester of the program. Since students take this course only in the first and second semesters during their first year, GPA was used for fourth year students and the participants were selected among those having a GPA of at least 3.00. These two criteria were set to gather rich data because some students were observed to have difficulty even in speaking, let alone using discourse markers, although they will all be teachers of English eventually. More specifically, five sample interviews were conducted with students who had various speaking scores that were lower than BA, and they provided hardly any data; they merely answered the interviewer's questions, and did not speak much, so the interview topics did not lead to a real conversation between the interviewer and the interviewee. The third criterion, i.e. study-abroad experience, was also used for fourth grade students, and not more than half of the participants in fourth year had a study-abroad experience (mostly within the Erasmus program). This specific criterion was set because many students in the department spend a semester at a European institution within the Erasmus program after their first year, and thus, have more opportunities to use English outside the classroom.

The interviews took around 20 minutes in average. The task characteristics to be followed in these interviews are identical to LOCNEC (Louvain Corpus of Native English Conversations), which are the second corpus to be used in the study. Each interview followed the same pattern: the interviewee was asked to talk about a travel experience, hobbies, a book, etc., which led to a conversation with the interviewer (i.e. the researcher); every interview ended with a short picture-based story-telling activity (Appendix II). The statistics regarding the corpus compiled in the scope of this study and information about the participants interviewed are presented in Table 3.3.

Table 3.3. *Descriptive statistics for the Turkish (TR) corpus*

	Duration of the Interview (mm:ss.ms)	Total No. of Words	No. of Words in Interviewee Speech	Year of Study	Age	Gender	Study Abroad
TR Participant 1	18:48.06	2,146	1,412	1	19	Female	-
TR Participant 2	16:21.07	1,711	1,140	1	19	Female	-
TR Participant 3	19:46.09	2,132	1,435	1	18	Female	-
TR Participant 4	15:04.06	1,925	1,277	1	19	Female	-
TR Participant 5	16:47.26	2,242	1,596	1	19	Male	-
TR Participant 6	13:21.14	1,542	872	1	19	Female	-
TR Participant 7	17:15.23	2,465	2,023	1	20	Female	Three Years
TR Participant 8	16:27.19	2,038	1,526	1	19	Male	-
TR Participant 9	12:08.15	1,273	743	1	18	Female	-
TR Participant 10	15:27.07	1,536	1,170	1	28	Male	-
TR Participant 11	13:54.19	1,645	1,159	1	18	Male	-
TR Participant 12	13:11.01	1,659	1,311	1	30	Female	-
TR Participant 13	15:20.15	1,842	1,118	1	19	Female	-
TR Participant 14	13:52.26	1,347	808	1	19	Female	-
TR Participant 15	15:49.07	2,318	1,471	1	18	Female	-
TR Participant 16	13:56.08	1,685	1,242	1	18	Female	-
TR Participant 17	12:05.13	1,092	528	1	18	Female	-
TR Participant 18	15:43.26	1,924	1,230	1	21	Female	-
TR Participant 19	10:43.12	1,204	813	1	19	Male	-
TR Participant 20	19:36.09	2,142	1,440	1	18	Female	-
TR Participant 21	13:44.19	1,655	1,164	1	18	Male	-
TR Participant 22	11:58.15	1,283	748	1	18	Female	-
TR Participant 23	15:39.07	2,328	1,478	1	18	Female	-
TR Participant 24	11:55.13	2,002	527	1	18	Female	-
TR Participant 25	16:11.07	1,721	1,145	1	19	Female	-
TR Participant 26	20:32.27	2,441	1,980	4	22	Male	-
TR Participant 27	20:14.22	2,148	1,760	4	21	Female	Six Months
TR Participant 28	13:36.09	1,267	803	4	21	Female	Six Months
TR Participant 29	16:07.28	2,163	1,778	4	22	Female	Six Months
TR Participant 30	13:51.27	1,476	752	4	22	Male	-

TR Participant 31	16:04.22	1,932	1,083	4	22	Male	Eight Months
TR Participant 32	14:56.03	1,492	944	4	22	Male	Six Months
TR Participant 33	11:24.16	1,301	711	4	21	Female	Six Months
TR Participant 34	14:41.15	1,668	1,029	4	21	Female	Six Months
TR Participant 35	16:35.23	2,072	1,699	4	21	Female	Eight Months
TR Participant 36	13:11.22	1,940	1,471	4	22	Female	Six Months
TR Participant 37	17:01.29	1,785	1,060	4	21	Female	-
TR Participant 38	14:26.01	1,390	900	4	21	Female	Six Months
TR Participant 39	15:20.20	1,867	1,497	4	24	Male	-
TR Participant 40	13:54.17	1,240	609	4	21	Female	-
TR Participant 41	13:34.10	1,774	1,134	4	21	Male	Six Years
TR Participant 42	14:31.13	1,743	978	4	22	Male	-
TR Participant 43	13:25.04	1,813	1,316	4	25	Female	-
TR Participant 44	13:19.16	1,323	616	4	21	Male	Six Months
TR Participant 45	16:51.29	1,795	1,065	4	21	Female	-
TR Participant 46	14:31.15	1,678	1,034	4	21	Female	Six Months
TR Participant 47	20:22.17	2,451	1,985	4	22	Male	-
TR Participant 48	14:22.13	1,753	983	4	22	Male	-
TR Participant 49	19:32.27	2,451	1,985	4	22	Male	-
TR Participant 50	13:22.10	1,784	1,139	4	21	Male	-
Whole Corpus	12:40:49	89,604	59,687				

The corpus of informal interview with Turkish university students thus contained a total of 89,604 words, and 59,687 words of interviewee speech. The interviews lasted 12 hours and 40 minutes in total, and around 15.2 minutes in average. The shortest interview was with TR Participant 19 and lasted 10 minutes 43 seconds, whereas the longest one was with TR Participant 26 and lasted 20 minutes 32 seconds. On the other hand, the interview with the most interviewee speech contained 2,023 words (TR Participant 7), while the one with the least interviewee speech had 527 words (TR Participant 24). The interviewee speech in the whole Turkish corpus was 1,194 words in average.

Among the 50 interviewees, 25 were in their first year and the remaining 25 in their last year studying English language teaching at a Turkish state university. As for their gender, 32 interviewees (64%) were female and 18 (36%) were male, which reflects the distribution within the program. The participants who contributed to the Turkish interview corpus were aged 20.58 years in average.

Of the first-year students in the Turkish corpus, only one student had a long-term experience abroad; TR Participant 12 spent nearly three years in the US and went to high school there. In the group of fourth-year students, 13 students had a long-term experience abroad; 12 of them studied at a European university for six months in the scope of the Erasmus program, and some also did internship within Erasmus, while one of the students, TR Participant 41 spent six year in Canada as the child of a Turkish immigrant family and went to primary school there.

3.2.2. Louvain corpus of native English conversations (LOCNEC)

The second corpus, LOCNEC (De Cock, 2004), was compiled within the scope of a project at the Centre for English Corpus Linguistics, Université Catholique de Louvain, and provided the native speaker reference data including informal interviews with 50 British university students majoring in English language and/or linguistics. The interviews were conducted at Lancaster University, United Kingdom. This corpus was purchased to be used in the analyses of this study. The statistics for the corpus are presented in Table 3.4 below.

Table 3.4. *Descriptive statistics for the native English (NS) speaker corpus (LOCNEC)*

	No. of Words	Interviewee Speech	Age	Gender
NS Participant 1	2,835	1,885	20	Female
NS Participant 2	3,128	2,088	19	Male
NS Participant 3	3,281	2,010	19	Female
NS Participant 4	3,161	2,130	18	Male
NS Participant 5	2,189	1,198	18	Male
NS Participant 6	3,338	2,081	19	Male
NS Participant 7	3,151	2,185	25	Male

NS Participant 8	3,635	3,176	21	Female
NS Participant 9	4,636	3,747	19	Female
NS Participant 10	3,377	2,409	20	Female
NS Participant 11	3,513	2,556	24	Female
NS Participant 12	3,465	2,268	24	Female
NS Participant 13	3,185	2,302	22	Female
NS Participant 14	2,930	2,045	20	Male
NS Participant 15	4,444	3,415	22	Female
NS Participant 16	3,393	2,361	23	Female
NS Participant 17	2,882	1,977	20	Male
NS Participant 18	3,940	3,055	19	Female
NS Participant 19	3,466	2,728	20	Male
NS Participant 20	2,216	1,542	19	Female
NS Participant 21	3,119	2,251	19	Male
NS Participant 22	4,433	3,834	25	Female
NS Participant 23	2,241	1,427	20	Female
NS Participant 24	3,073	1,935	19	Female
NS Participant 25	3,383	2,278	20	Female
NS Participant 26	5,095	4,261	24	Male
NS Participant 27	3,003	2,439	21	Male
NS Participant 28	2,349	1,138	18	Female
NS Participant 29	2,733	1,658	19	Female
NS Participant 30	3,006	2,085	22	Male
NS Participant 31	2,987	1,893	28	Female
NS Participant 32	2,930	2,059	21	Female
NS Participant 33	3,252	2,796	30	Female
NS Participant 34	3,118	2,598	22	Female
NS Participant 35	2,928	1,639	19	Female
NS Participant 36	2,964	1,877	25	Female
NS Participant 37	2,629	1,776	21	Male
NS Participant 38	3,631	2,934	26	Male
NS Participant 39	3,861	2,509	23	Female
NS Participant 40	2,802	1,746	22	Male
NS Participant 41	2,760	2,205	23	Male
NS Participant 42	2,537	1,993	27	Male

NS Participant 43	3,394	2,643	28	Male
NS Participant 44	3,407	2,652	19	Female
NS Participant 45	3,752	2,938	30	Female
NS Participant 46	3,451	3,141	20	Female
NS Participant 47	2,374	2,018	18	Female
NS Participant 48	2,815	2,335	22	Female
NS Participant 49	4,356	3,705	19	Male
NS Participant 50	3,185	2,554	19	Male
Whole Corpus	161,733	118,475		

The LOCNEC contained a total of 161,733 words including 118,475 words of interviewee speech. The interview with the most interviewee speech in the LOCNEC had 4,261 words (NS Participant 26) while the one with the least interviewee speech included 1,138 words (NS Participant 28). The average interviewee speech consisted of 2,370 words, and the average age of the participants was 21.6 years. Thirty of the participants were female and 20 were male.

3.2.3. Transcriptions of the interviews

As mentioned above, the Turkish interview corpus was compiled by following the same task characteristics employed in the LOCNEC. As for the transcriptions, the same transcription conventions were used as well (see Appendix III). A sample transcription of an interview for the Turkish corpus can be found in Appendix IV and a sample for the LOCNEC is provided in Appendix V. The transcriptions for the Turkish corpus were done by using a specific transcription tool named InqScribe. Figure 3.1 presents a screenshot of the transcriptions.

The turn of the interviewer was tagged with <A> * </A> while that of the interviewees was marked with <B> * </B>. In addition, specific tags such as for overlapping speech (</overlap>), filled and unfilled pauses (., .., …, er, erm), the use of a foreign word (<foreign>*</foreign>) were also employed during the transcription process. As can be seen in Figure 3.1, timestamps were also added to each turn so as to determine when the interviewee starts and ends talking in a turn. The transcriptions

were done by the researcher, and checked later for any spelling mistakes, or missing pauses and tags.

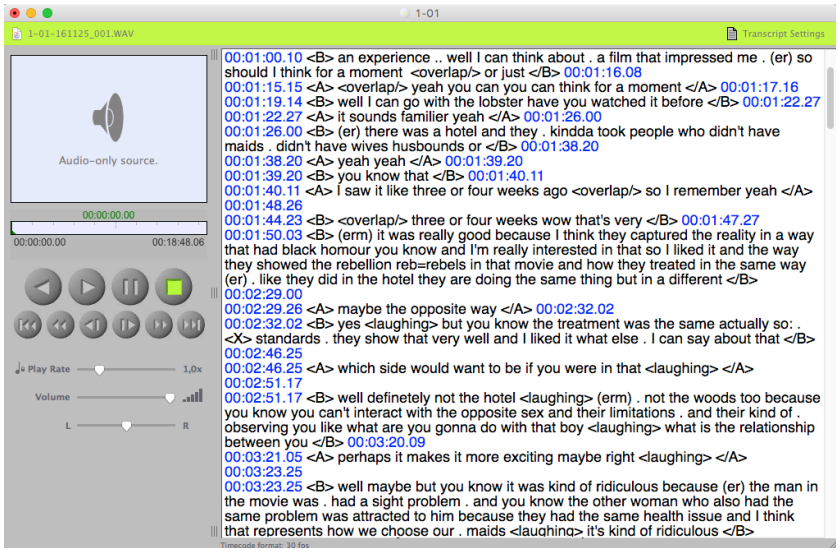


Figure 3.1. *Sample screenshot of a transcription file in InqScribe*

3.3. Data Analysis

3.3.1. Retrieval of discourse markers

The data analysis includes a quantitative and a qualitative part in accordance with the research questions. In the quantitative part, all instances of the selected discourse markers were retrieved using the Concord function of WordSmith Tools 6 (Scott, 2011) in the research corpora. Each instance was saved along with its frequency information and the broader context it has, and then, was manually checked to exclude any uses other than as a discourse marker. The raw frequencies were normalized to per 1,000 words to make meaningful comparisons with the frequencies of the discourse markers in the LOCNEC since the sizes of the two corpora may show differences.

Following the transcription of the interviews, all instances and frequencies of the discourse markers (i.e. *so, well, I mean, you know,* and *like*) were retrieved using WordSmith Tools 6 (Scott, 2011). Each instance was manually checked to exclude uses other than as a discourse marker. The instances of *so* as an adverb of degree or manner, in fixed expressions (e.g. *and so on*) and as a substitute (e.g. *I think so*) were excluded from the analysis. With regard to *well*, its use as an adverb (e.g. *well done*) and in fixed

expressions with the meaning of addition (e.g. *as well, as well as*) were discarded. About *you know*, the instances that are part of the question *do you know* and those that have a complement were not included. Lastly, the instances of *like* as a verb, preposition and conjunction (e.g. *like I said* replacing *as*) were also removed from the analysis. This manual check was also done by an independent researcher to ensure reliability.

After the frequencies were retrieved and the instances were manually checked to exclude other usages, the functions in which the discourse markers employed by the participants were determined. This was done by examining the concordance lines surrounding the instances. The functions reported in the literature were used to decide on for what function each instance of the discourse markers was employed.

3.3.2. Statistical comparison of the frequencies

In order to see the difference between the Turkish and British university students', and first and fourth year Turkish students' use of discourse markers, the log-likelihood statistic was used for a statistical comparison of the two corpora in terms of frequency through an online tool available at the website of Lancaster University Centre for Computer Corpus Research on Language (http://ucrel.lancs.ac.uk /llwizard.html). This test allows researchers to conduct tests for a significant difference in frequency between two corpora by using four simple figures:

1. The frequency of a word/expression in Corpus 1,
2. The total number of opportunities for that word/expression to occur in Corpus 1 (i.e. basically the total number of words it includes, or corpus size),
3. The frequency of a word/expression in Corpus 2,
4. The total number of opportunities for that word/expression to occur in Corpus 2 (i.e. basically the total number of words it includes, or corpus size).

The test produces a log-likelihood (LL) value, and the higher it is, the less likely it is that the result is a random one. Moreover, it should be above 3.84 for the difference to

be statistically significant at the level of p<.05 (i.e the 95% level). The advantage of the log-likelihood test is that it takes into account the size differences between the two corpora compared, which is quite important in corpus research. However, the log-likelihood test simply ignores individual differences between the texts that constitute a corpus, but takes the corpus as a whole. In such a case, if, say, half of the texts do not include the lexical item under examination, but the other half include it with a high frequency, then there may not be a significant difference between two corpora in the log-likelihood test. This can be important when the comparison of two corpora actually aims to compare two groups of learners. In this respect, in addition to the log-likelihood statistic, a further inferential test was also used to address this issue.

In inferential statistics, there are parametric (e.g. independent samples t-test or ANOVA) and non-parametric tests (e.g. Mann-Whitney U or Kruskal–Wallis) which are chosen based on whether the data are normally distributed or not. For this, the Shapiro-Wilks test was run in SPSS, and the result revealed a significant difference (p<.05) for all discourse markers examined for both groups, which means that the data did not show a normal distribution. This can be reasonable in corpus research because the data include frequencies, and some of the texts may have a high frequency of an item while other may have zero frequency. As a result, since the data did not show normal distribution, a non-parametric test needed to be used, and because two groups/corpora were compared, the Mann-Whitney U test, which is known as the non-parametric alternative of the widely known Independent Samples t-test, was also used in the present study.

In the present study, the frequencies of the five discourse markers in the Turkish corpus and LOCNEC were used in the log-likelihood and Mann-Whitney U tests. In this way, it could be revealed whether the frequencies of the discourse markers in the two corpora showed statistically significant differences.

3.3.3. Functional analysis of discourse markers

In the qualitative part, each instance of the discourse markers was examined to determine its function in discourse. This process included a total of 4661 instances of the five discourse markers under examination; 1093 instances in the Turkish corpus and

3568 in the LOCNEC. Based on the concordance lines, and when necessary the previous and following turns, all instances were checked one by one along with the immediate context they occurred. When deciding on what function an instance serves in discourse, the following steps were followed:

- The linguistic surrounding of the instance as well as the context and the topic being talked about were considered,
- Previous and following utterances of the same turn that go beyond the brief concordance lines provided by WordSmith Tools were checked and put into consideration,
- Previous and following turns were checked where a function possibly makes reference to other turns,
- The functions reported/documented in studies presented in Table 3.2 above were also taken as reference, but the researcher did not necessarily try to apply them to the data.

To sum up, a data-driven approach was adopted; rather than trying to fit the data into a predefined typology or classification, a bottom-up approach was employed and the functions appearing in the data were revealed. Since there is no generally agreed functional typology or categorisation regarding discourse markers, the functions detected in the research corpora were presented according to the domains, i.e. textual, interpersonal and interactional, in which they served. As elaborated in the literature review above, textual functions occur with regard to the organisation of discourse by signalling how different parts of the speech are related to each other (Aijmer and Simon-Vandenbergen, 2011). However, interpersonal functions operate in a more subjective level (Brinton, 1996), and help the speaker share how he/she feels by conveying an attitude or evaluation (Aijmer 2002). As for interactional functions, they are related to the creation of a sphere of sharing (Brinton, 1996), and help the speaker to hint to the hearer to interact with him/her to create common knowledge (Müller, 2005). In brief, the functions detected in the corpora examined in this study were presented and analysed based on the domains in which they operated.

Another issue to which due attention was paid was the reliability of the functional analysis. For this, an independent researcher who is a native speaker of English examined a part of the instances of the five discourse markers after the researcher himself conducted the functional analysis. Accordingly, 10% of the interviews in the Turkish corpus and LOCNEC were coded by another researcher who is a lecturer at a Turkish university, and any disagreements were resolved through discussion. The agreement rate was initially found to be about 75%, while the sufficient rate for reliable results is generally accepted as 70% (Müller, 2005, p. 28; Mei, 2012, p. 55).

Finally, after the functions were determined for each instance of the five discourse markers, their frequencies and percentages were calculated and presented individually, and under the three domains described above. Then, in order to compare the number of instances extracted from the two corpora in terms of statistical significance, the two-sample frequency comparison utility of the project "Statistical Inference: A Gentle Introduction for Linguists (SIGIL)" (http://sigil.collocations.de/wizard.html) was used. This utility works based on chi-square (X2) and log-likelihood statistics (G2), and chooses between these tests based on the size of the data. It gives the statistically significant proportional differences between the functions reported in the two corpora.

3.3.4. Evaluation of oral fluency

As is stated above, a number of variables related to utterance such as speech rate, articulation rate, mean length of runs, and number and mean length of pauses per minute, which contribute to utterance fluency, are used to evaluate overall fluency of oral performance along with the perceived ratings of native speakers, which is evaluated as perceived fluency (Kormos and Denes, 2004). In the present study, both temporal variables (i.e. utterance fluency) were calculated and a perceived rating for each interviewee participated in the Turkish corpus was done to have an overall picture of the interviewees' speech fluency. For the temporal variables, the following steps were followed:

- Where every single turn started and ended in all interviews was coded with a timestamp (e.g. 00:12:01.15 <B> I guess it's . because it's not original . </B> 00:12:06.00),
- The timestamps were used to calculate the total time used by the interviewees in their turns,
- The parts of the interviews during which only the interviewees spoke were trimmed from the interview recordings,
- The trimmed audio files including only the interviewee speech were analysed by using Praat (Boersma and Weenink, 2007), which is a free computer software package for the scientific analysis of speech in phonetics,
- By means of a Praat script (de Jong and Wempe, 2008), the number and duration of pauses/silences as well as the phonation time were calculated in each interviewee's speech (see Figure 3.2 below),
- As a result, the following variables were calculated for utterance fluency:
    - speech rate in syllables per minute (SPM) was obtained by dividing the total number of syllables by total time,
    - speech rate in words per minute (WPM) was obtained by dividing the total number of words by total time,
    - articulation rate was obtained by dividing the total number of syllables by phonation time,
    - mean length of runs was obtained by dividing the total number of syllables by the total number of pauses,
    - mean pause duration was obtained by dividing the total duration of pauses by the total number of pauses,
    - mean duration of syllables was obtained by dividing the phonation time by the total number of syllables.

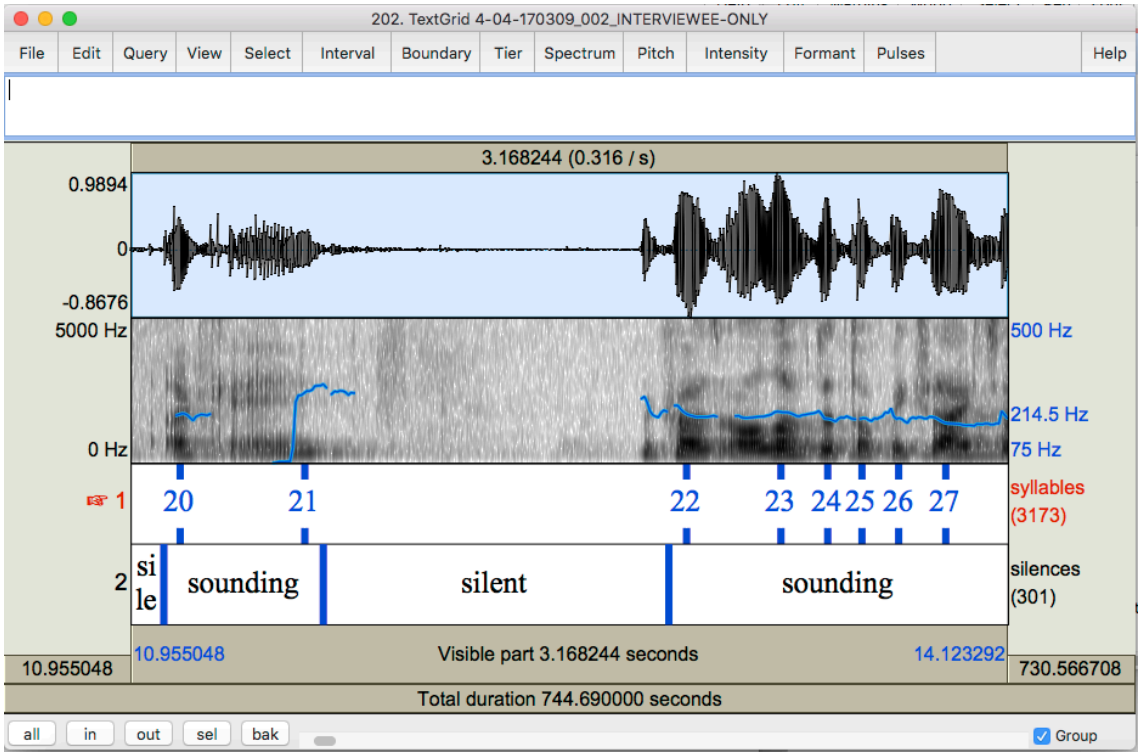


Figure 3.2. *Sample screenshot of a Praat window taken during the analysis*

After the above-mentioned temporal variables were calculated, Turkish interviewees' interview recordings were listened to by a native English speaker for perceived fluency. The recordings were rated on a scale from 1 being extremely dysfluent to 9 being extremely fluent by a native English speaker working as a lecturer at a Turkish university. She was instructed to evaluate each interviewee in terms of his/her flow and smoothness of speech, and was not told that the primary focus of this study was the use of discourse markers.

Finally, the descriptive statistics were presented and reported with regard to the interviewees' utterance fluency and perceived fluency. Then, a Pearson correlation analysis was conducted to examine the relationship between the Turkish interviewees' use of the discourse markers, and their utterance and perceived fluency.

3.4. Pilot Study

Prior to the collection and analysis of the research data described above, a pilot study was conducted with a smaller sample to have a preliminary insight about the extent and variety of discourse marker use in the target population, as well as about how

the analysis would be done. This pilot study included only the first five participants interviewed for the Turkish corpus described above.

3.4.1. Findings of the pilot study

The frequencies of the discourse markers occurred in the pilot corpus were normalised to per 100 words since the total number of words produced by the participants were different. On the other hand, the findings regarding the functions were supported with sample quotations from the interviews.

3.4.2. Frequencies of discourse markers

In the pilot study, the five discourse markers *so, well, I mean, you know,* and *like* occurred a total of 151 times (2.20 times per 100 words) in the pilot corpus. Among these discourse markers, *so* was the most frequent with a raw frequency of 79 times (1,15 times per 100 words). In other words, *so* constituted about 52% of all instances of these discourse markers. Table 3.5 presents the frequencies of these discourse markers.

Table 3.5. *Frequencies of the discourse markers in the pilot corpus*

	Raw Frequency	Normalised Frequency	%
I mean	5	0.07	3.30
well	13	0.19	8.60
you know	23	0.33	15.23
like	31	0.45	20.52
so	79	1.15	52.31
Total	151	2.20	100

You know (0.33 times per 100 words) and *I mean* (0.45 times per 100 words) came after *so* while the least frequent of the discourse markers was *I mean* that occurred 0.07 times per 100 words, which was followed by *well* that occurred 0.19 times per 100 words. In addition to the frequencies of these discourse markers in overall, the use by each participant should also be examined because a single participant may employ an excessive number of a certain discourse marker, which can inflate the findings. Figure

3.3 below shows the distribution of the discourse markers used by the participants in the pilot corpus.

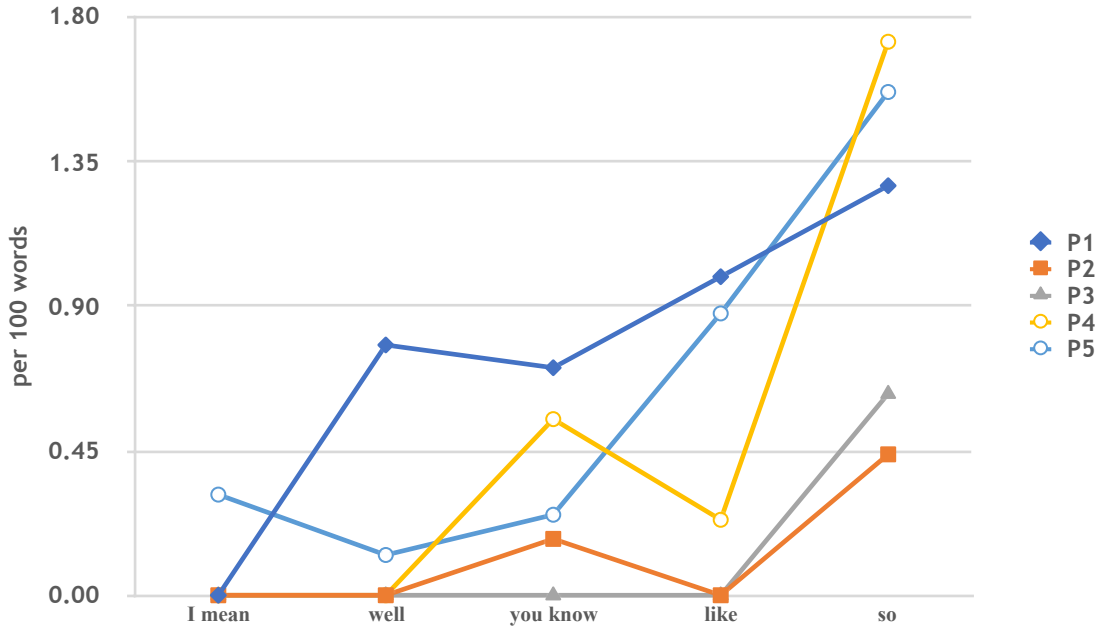


Figure 3.3. *Distribution of the discourse markers across the participants*

The only participant that used all the discourse markers was Participant 5 at different frequency levels, while four out of the five discourse markers were not used by Participant 3 who only employed *so* in her speech. Participant 1 used *well*, *you know* and *like* more frequently than the rest of the participants, but did not employ *I mean*. Participant 2 made use of only *you know* and *so*, whereas Participant 4 did not employ *I mean* and *well*, but used *you know*, *like* and *so* in that the frequency of *so* was the highest in her speech compared to the other participants.

The native speaker reference corpus (i.e. LOCNEC) was not yet available to the researcher at the time of the pilot study to make a comparison. However, the frequency data reported in Buysse (2010) who selected and analysed 20 of the 50 interviews in the LOCNEC corpus were referred to. Figure 3.4 compares the frequency of the five discourse markers by Turkish students as retrieved in this pilot study, and British university students as reported in Buysse (2010).

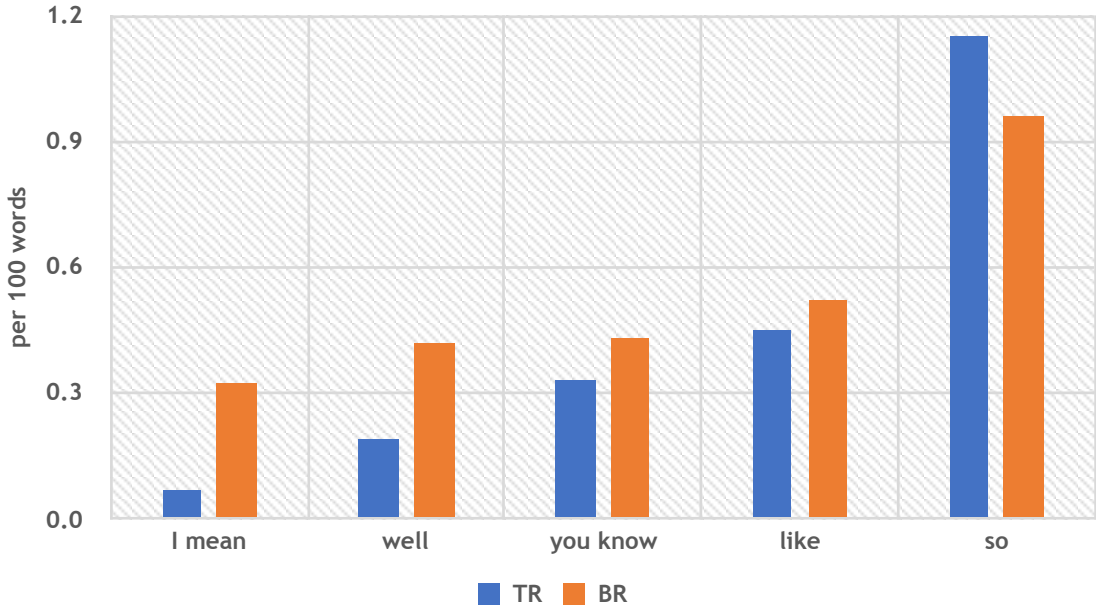


Figure 3.4. *Comparison of the discourse markers used by Turkish and British students*

Above all, it should be noted that this comparison was based on data coming from five Turkish students and 20 British students, and thus, it does not seem plausible to draw a conclusion. However, an overall insight might be gained through these normalised frequencies. In this respect, there seems to be a gradual increase in the frequencies from *I mean* as the least frequent to *so* as the most frequent in both the Turkish students' and British students' speech. Four of the five discourse markers, *I mean, well, you know* and *like* were used more frequently by the British students, whereas *so* was overused by the Turkish students. The difference in-between is the closest in the use of *like*. Yet, whether these differences reveal a statistical significance should also be analysed, but it is not reasonable to make such an analysis without having the actual native speaker data. On the other hand, the differences should also be evaluated in terms of the functions in which the individuals used the discourse markers. This is because, for example, *so* was overused by the Turkish students who may have employed it only in a single function, whereas the use of their native counterparts may include a variety of functions which never occur in the Turkish students' speech.

3.4.3. Functions of discourse markers in the pilot study

3.4.3.1. So

The most frequent of the five discourse markers in the pilot study, *so*, was used by the participants in six different functions, which can be argued to have a considerable variety. In other words, they did not employ *so* in one or two textual functions, but also used it in various interpersonal functions. About 57% of the 79 instances of *so* as a discourse markers was employed to mark/indicate a result by the participants:

(1) <A> I think there was some sort of punishment right </A>
<B> before that the woman also had the same problem but she didn't use glasses she just couldn't see very well . and the man had glasses *so* they were kind of they connected over that . (…) </B> (Participant 1)

Here, *so* indicates the result of a situation, in that Participant 1 says that the two characters in the movie got close because they had a common characteristic. The second most common function of *so* was marking the main idea which constituted 15 instances (19%):

(2) <A> do you usually go . such holidays with your family </A>
<B> no we usually go to Sivas <laughing> . but . I like Sivas because my relatives live there . and we are really close we love each other very much we miss each other . *so* it's good to go to Sivas as well </B> (Participant 3)

Within this conversation, Participant 3 says that she likes her hometown, explains why and then she lastly draws the interlocutor's attention to the main idea that her hometown is good to visit and she likes it. Other functions of *so* that the participants employed were mark an implied result/yield the floor (11%), draw a conclusion (4%), discourse boundary/hold the floor (4%), and mark a question (4%).

3.4.3.2. Well

Eleven of the 13 instances of *well* as a discourse marker was employed by Participant 1, and the remaining two instances by Participant 5. These two participants mostly used this discourse marker to mark an answer to a question (46%):

(3) <A> how did you manage to choose such a movie </A>
<B> *well* I like those <starts laughing> kind of movies <stops laughing> (erm) </B> (Participant 1)

The rest of the instances of *well* was employed by these two participants with the function of contributing an opinion, evaluating a previous statement, introducing the next idea/scene, searching for the right phrase, moving to the main story/idea, and continuing an opinion/answer. Although this discourse marker seems to be used for different functions, most of these occurred in the speech of only Participant 1.

3.4.3.3. I mean

I mean occurred only five times in the pilot corpus, and was only employed by Participant 5. In two of these instances, he used *I mean* to express an evaluation, for example:

(4) <A> okay so what happens (er) at last </A>
<B> *I mean* I guess it's her friends I'm not sure . could be a museum but . I don't know </B> (Participant 5)

In this quotation, the participant seems to convey his evaluation regarding how the story in the picture description task ends. The rest of the functions he employed included drawing a conclusion, exemplification and mark a quotation.

3.4.3.4. You know

You know was the third most frequent among the discourse markers examined in the study with a frequency of 23 times. It was mostly used with the function of referring to shared knowledge (38%):

(5) <A> you hated it </A>
<B> yes . just what I mentioned *you know* the circle and it <starts laughing> just doesn't go anywhere <stops laughing> and I don't like it </B> (Participant 1)

In this quotation, Participant 1 refers to a comment she made about the TV series she watched before. Another common function employed by the participants was providing justification:

(6) <A> which side would want to be if you were in that <laughing> </A>
<B> well definetely not the hotel <laughing> (erm) . not the woods too because *you know* you can't interact with the opposite sex and their limitations . and their kind of . observing you like what are you gonna do with that boy <laughing> what is the relationship between you </B> (Participant 2)

Here, Participant 2 provides a justification for her answer to the interviewer's question by using *you know*. Other functions for which the participants employed this discourse marker included seeking confirmation, correcting hearer assumption and quoting.

3.4.3.5. Like

Like was the second most frequent discourse marker employed by the participants. The most common function for which it was employed was marking an approximate number or quantity (29%):

(7) <A> when was this </A>
<B> (er) I went there *like* five years ago . (er) it was definitely an amazing trip for me </B> (Participant 5)

In this extract, Participant 5 is not sure about the time he wants to indicate and uses *like* to approximate it. Among the other functions, *like* was also frequently employed to mark quotations (25%):

(8) <A> which side would want to be if you were in that <laughing> </A> 00:02:51.17
<B> well definetely not the hotel <laughing> (erm) . not the woods too because you know you can't interact with the opposite sex and their limitations . and their kind of . observing you *like* what are you gonna do with that boy <laughing> (Participant 1)

In this use, Participant 1 produces an utterance in the form a quotation which is marked by *like*. Other functions for which this discourse marker was used included searching for an appropriate expression and introducing an example.

3.4.4. Speaking fluency and discourse marker use in the pilot study

Native-like use of discourse markers is argued to be crucial for learners to express themselves in the target language in a fluent way (Hellermann and Vergun, 2007; Sankoff et al., 1997). However, whether the use of discourse markers is a characteristic of fluent language use has not been addressed in the literature before. In this pilot study, the participants' speech rates in the interviews were calculated through the total number of syllables divided by the amount of total time required to produce the speech sample. Speech rate was taken as a measure of fluency because it is reported to be a temporal variable that strongly correlate with the fluency ratings of native speakers (Kormos and Denes, 2004). Furthermore, the recordings of the interviews were rated on a scale from 1 being extremely dysfluent to 9 being extremely fluent by a native English speaker working as a lecturer at a university in Istanbul, Turkey. Table 3.6 presents the overall findings regarding the variables of speech rate and native speaker rating.

Table 3.6. *Participants' speech rate and fluency ratings*

	Total No. of Syllables	Total Time (seconds)	Speech Rate (Average no. of syllables per minute)	NS Fluency Rating 1-9
Participant 1	1,814	747.62	145.58	6
Participant 2	1,533	677.76	135.71	3
Participant 3	1,939	747.23	155.69	4
Participant 4	1,611	539.49	179.16	5
Participant 5	2,240	681.24	197.28	7

NS: Native Speaker

Overall findings show that the participant with the lowest speech rate were rated by the native speaker rater as the least fluent, whereas the one with the highest speech rate as the most fluent among the five participants. With the participants' speech rate

and native speaker fluency rating as variables of speaking fluency at hand, a correlation analysis was carried out regarding the relationship between these variables and discourse marker use. The results of this analysis are presented in Table 3.7 below in the form of a correlation matrix.

Table 3.7. *Relationship between discourse marker use, and speech rate and fluency rating*

	Use of five DMs	Speech rate	NS fluency rating
Use of five DMs	1	.439	.894*
Speech rate	.439	1	.687
NS fluency rating	.894*	.687	1

DMs: Discourse markers, NS: Native speaker

*Correlation is significant at the 0.05 level (2-tailed).

The result showed that there was a moderate positive correlation (r=.43) between the total frequency of the five discourse markers employed by the participants and their speech rates. Besides, there was a strong and significant correlation (r=.89, p<.05) between the total frequency of the five discourse markers and the participants' fluency as rated by an English native speaker. These two findings show that discourse marker use can be related to, or a characteristic of fluent speech. However, it would be more meaningful to evaluate the result of this analysis with a larger sample, when the data gathering process would be completed. This is because the pilot sample included only five students although more is always regarded as better in statistical analyses such as correlation. Furthermore, other temporal variables such as mean length of runs and silent pauses in addition to speech rate could provide more insights about how the employment of discourse markers relate to speaking fluency.

After a pilot study was conducted to have a preliminary insight about the extent and variety of discourse marker use in the target population, as well as about how the analysis would be done, the actual data collection and analysis procedure were implemented, and the results are presented in the following section.

4. RESULTS

4.1. Initial Findings on Discourse Markers in the Research Corpora

After the Turkish corpus was compiled and the LOCNEC was obtained, the five discourse markers, namely *so, well, you know, I mean* and *like,* were searched in the both corpora. All instances and frequencies were retrieved using WordSmith Tools 6 (Scott, 2011), and each instance was manually checked to exclude uses other than as a discourse marker. The instances of *so* as an adverb of degree or manner, in fixed expressions (e.g. *and so on*) and as a substitute (e.g. *I think so*) were excluded from the analysis. With regard to *well*, its use as an adverb (e.g. *well done*) and in fixed expressions with the meaning of addition (e.g. *as well, as well as*) were discarded. About *you know*, the instances that are part of the question *do you know* and those that have a complement were not included. Lastly, the instances of *like* as a verb, preposition and conjunction (e.g. *like I said* replacing *as*) were also removed from the analysis. Table 4.1 presents a summary of all instances retrieved and the number of remaining instances that were discourse markers.

Table 4.1. *Raw frequencies of all instances and discourse markers*

	Turkish Corpus		LOCNEC	
	All	DM	All	DM
so	561	404	1,628	1,364
like	753	298	1,347	633
you know	260	250	633	610
I mean	96	87	443	430
well	85	54	754	531
Total	1,746	1,093	4,805	3,568

DM: Discourse marker

In total, 6,551 instances in both corpora including 1,746 in the Turkish corpus and 4,806 in the LOCNEC were retrieved through WordSmith Tools, and manually checked by examining the larger contexts when necessary. At the end, a total of 4,661 instances in both corpora including 1,093 in the Turkish corpus and 3,568 in the LOCNEC were

found to be used as discourse markers. As a discourse marker, *so* occurred 404 times in the Turkish corpus and 1,364 times in the LOCNEC, whereas *like* occurred 298 times in the Turkish corpus, and 633 times in the LOCNEC. Besides, *you know*, *I mean* and *well* respectively occurred 250, 87 and 54 times in the Turkish corpus, and 610, 430 and 531 times in the LOCNEC.

4.2. Frequencies of Discourse Markers in the Turkish Corpus

After the non-discourse-marker instances were retrieved and the raw frequencies of the five discourse markers, namely *so, well, you know, I mean* and *like,* were obtained, the normalised frequencies were calculated as per 1,000 words since the two corpora to be compared were not of the exact same size. The raw and normalised frequencies of these discourse markers that occurred in the Turkish corpus are presented in Table 4.2 below.

Table 4.2. *Frequencies of the discourse markers in the Turkish corpus*

	Raw Frequency	Normalised Frequency	%
well	54	0.90	4.92
I mean	87	1.45	7.93
you know	250	4.18	22.86
like	298	4.99	27.29
so	404	6.76	36.98
Total	1,093	18.28	100

In the Turkish corpus, the five discourse markers occurred a total of 1,093 times (18.28 times per 1,000 words). Among these discourse markers, *so* was the most frequent with a raw frequency of 404 times (6,76 times per 1,000 words). In other words, *so* constituted about 36% of all instances of these discourse markers. It was followed by *like* that occurred 298 times (4.99 times per 1,000 words), *you know* with an occurrence of 250 times (4.18 times per 1,000 words) and *I mean* with 87 instances (1.45 times per 1,000 words). The least frequent of these five discourse markers was

well that occurred 54 times (0.90 times per 1,000 words) in the Turkish corpus. The frequency variations can be represented in Figure 4.1.

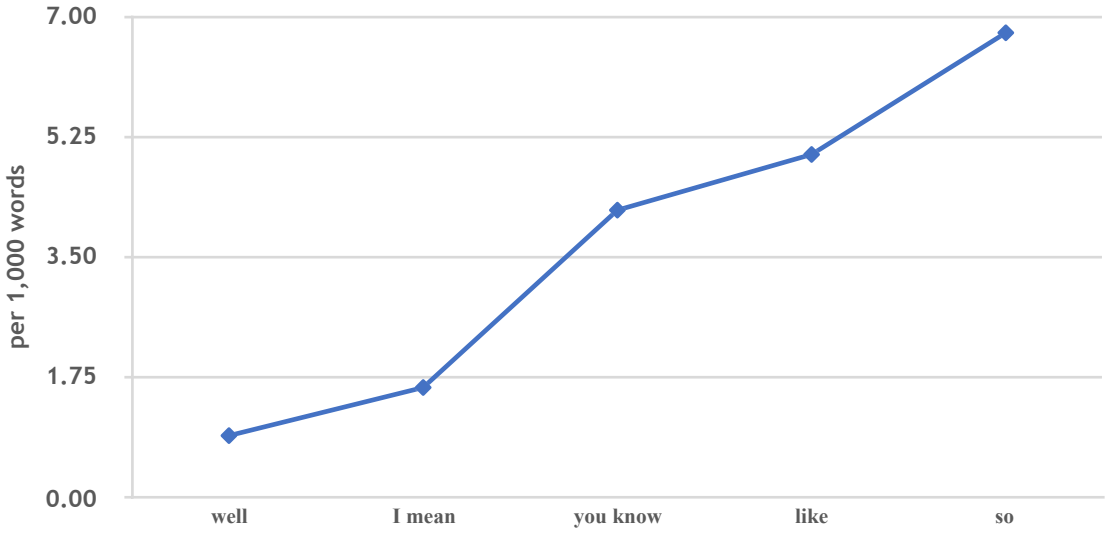


Figure 4.1. *Frequency variation of the discourse markers in the Turkish corpus*

This result largely overlaps with the pilot analysis in that the frequency of *so* is considerably higher than those of *well* and *I mean* in the whole corpus. Although the frequencies of the discourse markers seem to vary across the corpus, the variation among the individual participants is also quite diverse. The frequencies of the discourse markers employed by each participant in the Turkish corpus can be found in Table 4.3 below.

Table 4.3. *Frequencies of discourse markers employed by participants in the Turkish corpus*

		so	like	you know	I mean	well			so	like	you know	I mean	well
TR P1	R	18.00	10.00	10.00	0.00	11.00	TR P26	R	1.00	4.00	1.00	0.00	0.00
	N	12.75	7.08	7.08	0.00	7.79		N	0.51	2.02	0.51	0.00	0.00
TR P2	R	5.00	0.00	2.00	0.00	0.00	TR P27	R	23.00	0.00	2.00	0.00	1.00
	N	2.92	0.00	1.17	0.00	0.00		N	13.07	0.00	1.14	0.00	0.57
TR P3	R	8.00	1.00	0.00	0.00	1.00	TR P28	R	20.00	0.00	0.00	0.00	0.00
	N	5.57	0.70	0.00	0.00	0.70		N	24.91	0.00	0.00	0.00	0.00
TR P4	R	22.00	4.00	7.00	0.00	0.00	TR P29	R	14.00	2.00	0.00	33.00	0.00
	N	17.23	3.13	5.48	0.00	0.00		N	7.87	1.12	0.00	18.56	0.00

TR P5	R	25.00	14.00	4.00	5.00	3.00	TR P30	R	2.00	2.00	0.00	0.00	0.00
	N	15.66	8.77	2.51	3.13	1.88		N	2.66	2.66	0.00	0.00	0.00
TR P6	R	7.00	11.00	0.00	0.00	0.00	TR P31	R	0.00	11.00	15.00	0.00	17.00
	N	8.03	12.61	0.00	0.00	0.00		N	0.00	10.16	13.85	0.00	15.70
TR P7	R	26.00	67.00	4.00	3.00	0.00	TR P32	R	0.00	3.00	0.00	0.00	0.00
	N	12.85	33.12	1.98	1.48	0.00		N	0.00	3.18	0.00	0.00	0.00
TR P8	R	9.00	1.00	0.00	32.00	13.00	TR P33	R	2.00	4.00	11.00	0.00	0.00
	N	5.90	0.66	0.00	20.97	8.52		N	2.81	5.63	15.47	0.00	0.00
TR P9	R	3.00	1.00	0.00	4.00	0.00	TR P34	R	1.00	1.00	1.00	1.00	0.00
	N	4.04	1.35	0.00	5.38	0.00		N	0.97	0.97	0.97	0.97	0.00
TR P10	R	3.00	9.00	6.00	1.00	0.00	TR P35	R	14.00	22.00	1.00	2.00	0.00
	N	2.56	7.69	5.13	0.85	0.00		N	8.24	12.95	0.59	1.18	0.00
TR P11	R	0.00	2.00	1.00	0.00	0.00	TR P36	R	18.00	3.00	19.00	0.00	0.00
	N	0.00	1.73	0.86	0.00	0.00		N	12.24	2.04	12.92	0.00	0.00
TR P12	R	26.00	8.00	36.00	0.00	0.00	TR P37	R	4.00	2.00	0.00	0.00	0.00
	N	19.83	6.10	27.46	0.00	0.00		N	3.77	1.89	0.00	0.00	0.00
TR P13	R	9.00	2.00	18.00	0.00	1.00	TR P38	R	8.00	0.00	0.00	0.00	0.00
	N	8.05	1.79	16.10	0.00	0.89		N	8.89	0.00	0.00	0.00	0.00
TR P14	R	4.00	2.00	0.00	0.00	0.00	TR P39	R	6.00	0.00	19.00	0.00	0.00
	N	4.95	2.48	0.00	0.00	0.00		N	4.01	0.00	12.69	0.00	0.00
TR P15	R	8.00	0.00	6.00	4.00	1.00	TR P40	R	5.00	0.00	0.00	0.00	0.00
	N	5.44	0.00	4.08	2.72	0.68		N	8.21	0.00	0.00	0.00	0.00
TR P16	R	25.00	26.00	42.00	1.00	0.00	TR P41	R	5.00	11.00	10.00	0.00	2.00
	N	20.13	20.93	33.82	0.81	0.00		N	4.41	9.70	8.82	0.00	1.76
TR P17	R	2.00	0.00	0.00	0.00	0.00	TR P42	R	13.00	3.00	1.00	0.00	0.00
	N	3.79	0.00	0.00	0.00	0.00		N	13.29	3.07	1.02	0.00	0.00
TR P18	R	13.00	29.00	2.00	1.00	0.00	TR P43	R	0.00	1.00	3.00	0.00	0.00
	N	10.57	23.58	1.63	0.81	0.00		N	0.00	0.76	2.28	0.00	0.00
TR P19	R	2.00	12.00	1.00	0.00	0.00	TR P44	R	2.00	1.00	5.00	0.00	0.00
	N	2.46	14.76	1.23	0.00	0.00		N	3.25	1.62	8.12	0.00	0.00
TR P20	R	8.00	1.00	0.00	0.00	1.00	TR P45	R	4.00	2.00	0.00	0.00	0.00
	N	5.56	0.69	0.00	0.00	0.69		N	3.76	1.88	0.00	0.00	0.00
TR P21	R	0.00	2.00	1.00	0.00	0.00	TR P46	R	1.00	1.00	1.00	1.00	0.00
	N	0.00	1.72	0.86	0.00	0.00		N	0.97	0.97	0.97	0.97	0.00
TR P22	R	3.00	1.00	0.00	4.00	0.00	TR P47	R	1.00	4.00	1.00	0.00	0.00
	N	4.01	1.34	0.00	5.35	0.00		N	0.50	2.02	0.50	0.00	0.00

TR P23	R	8.00	0.00	6.00	4.00	1.00	TR P48	R	13.00	3.00	1.00	0.00	0.00
	N	5.41	0.00	4.06	2.71	0.68		N	13.22	3.05	1.02	0.00	0.00
TR P24	R	2.00	0.00	0.00	0.00	0.00	TR P49	R	1.00	4.00	1.00	0.00	0.00
	N	3.80	0.00	0.00	0.00	0.00		N	0.50	2.02	0.50	0.00	0.00
TR P25	R	5.00	0.00	2.00	0.00	0.00	TR P50	R	5.00	11.00	10.00	0.00	2.00
	N	4.37	0.00	1.75	0.00	0.00		N	4.39	9.66	8.78	0.00	1.76

TR P#: Turkish Participant, R: Raw frequency, N: Normalised Frequency

To start with *so*, five participants in the Turkish corpus (TR P11, TR P21, TR P31, TR P32, TR P43) never employed this discourse marker in their interviews, whereas two participants (TR P28 and TR P16) with the highest frequency employed it 24.91 and 20.13 times per 1,000 words, respectively. The frequencies of the remaining 43 participants ranged from 19.83 to 0.50 times per 1,000 words.

As for *like*, 11 participants (TR P2, TR P15, TR P17, TR P23, TR P24, TR P25, TR P27, TR P28, TR P38, TR P39 and TR P40) never used it as a discourse marker, while TR P7 and TR P18 were the two participants who employed it the most with the normalised frequencies of 33.12 and 23.58 times per 1,000 words. The frequencies of the remaining 37 participants were between 20.93 and 0.66 times per 1,000 words.

With regard to *you know*, it did not occur in the speech of 17 participants, whereas it was used by TR P16 the most with a frequency of 33.82 times per 1,000 words. The frequencies of the rest of the participants ranged between 27.46 and 0.50 times per 1,000 words.

Regarding *I mean*, 36 participants never actually used it in their speech, while TR P8 was the participant that employed it the most with a frequency of 20.97 times 1,000 words. The frequencies of the remaining 13 participants were between 18.56 and 0.81 times per 1,000 words.

Lastly, the least frequently occurred discourse marker, *well*, did not occur in the speech of 38 participants during the interviews, whereas TR P31 employed it 15.70 times per 1,000 words, which is the highest frequency for this discourse marker. The frequencies of the remaining 11 participants varied from 8.52 to 0.57 times per 1,000 words.

To better demonstrate the variation in the frequencies of the discourse markers employed by each individual, their performance profiles are presented in Figure 4.2.

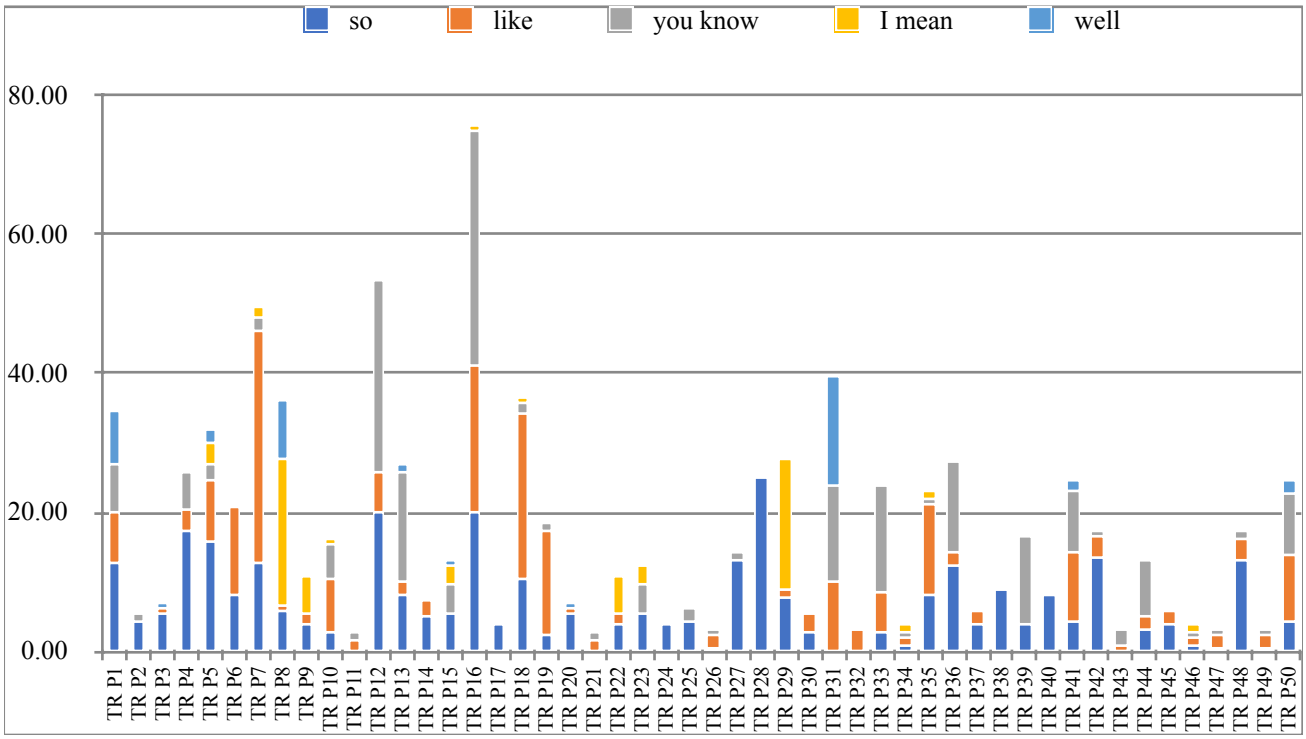


Figure 4.2. *Performance profiles of the Turkish students for the five discourse markers (in per 1,000 words) in the Turkish corpus*

As is seen in Figure 4.2, among the five discourse markers, *so* had the widest coverage within the corpus in that 45 out of 50 participants used it in their interviews. On the other hand, *well* was employed by the least number of participants; only 12 out of 50 participants employed it during the interviews. Furthermore, *like* did not occur in the speech of 11 participants, and nine participants used it only once.

4.3. Frequencies of Discourse Markers in the LOCNEC

The raw and normalised frequencies of the five discourse markers, namely *so, well, you know, I mean* and *like,* were also obtained as per 1,000 words for the LOCNEC. The raw and normalised frequencies of these discourse markers that occurred in the LOCNEC are presented in Table 4.4 below.

Table 4.4. *Frequencies of the discourse markers in the LOCNEC*

	Raw Frequency	Normalised Frequency	%
I mean	430	3.62	12.02
well	531	4.48	14.88
you know	610	5.14	17.07
like	633	5.34	17.74
so	1,364	11.52	38.27
Total	3,568	30.10	100

In the LOCNEC, the five discourse markers occurred a total of 3,568 times (30.10 times per 1,000 words). Among these discourse markers, *so* was the most frequent with a raw frequency of 1,364 times (11.52 times per 1,000 words). In other words, *so* constituted about 38% of all instances of these discourse markers in the LOCNEC. It was followed by *like* that occurred 633 times (5.34 times per 1,000 words), *you know* with an occurrence of 610 times (5.14 times per 1,000 words) and *well* with 531 instances (4.48 times per 1,000 words). The least frequent of these five discourse markers was *I mean* that occurred 430 times (3.62 times per 1,000 words). The frequency variations in the LOCNEC can be represented in Figure 4.3.

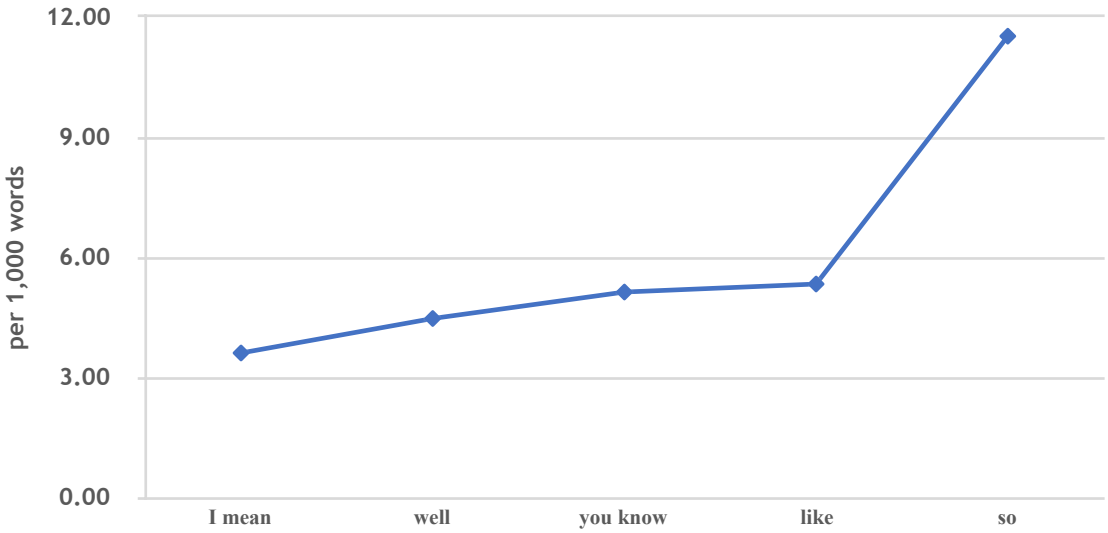


Figure 4.3. *Frequency variation of the discourse markers in the LOCNEC*

Among the five discourse markers, *so* was used with a considerably higher frequency than those of the remaining in the LOCNEC. In overall, the frequencies seem to vary across *I mean*, *well*, *you know* and *like* with quite small differences. As for variation among the individual participants, the frequencies of the discourse markers employed by each participant in the LOCNEC are presented in Table 4.5 below.

Table 4.5. *Frequencies of discourse markers employed by participants in the LOCNEC*

		so	like	you know	I mean	well			so	like	you know	I mean	well
NS P1	R	14.00	2.00	7.00	0.00	12.00	NS P26	R	23.00	4.00	25.00	31.00	11.00
	N	7.43	1.06	3.71	0.00	6.37		N	5.40	0.94	5.87	7.28	2.58
NS P2	R	16.00	34.00	4.00	2.00	25.00	NS P27	R	27.00	14.00	0.00	6.00	18.00
	N	7.66	16.28	1.92	0.96	11.97		N	11.07	5.74	0.00	2.46	7.38
NS P3	R	30.00	18.00	11.00	2.00	8.00	NS P28	R	13.00	2.00	12.00	9.00	9.00
	N	14.93	8.96	5.47	1.00	3.98		N	11.42	1.76	10.54	7.91	7.91
NS P4	R	17.00	3.00	5.00	5.00	17.00	NS P29	R	13.00	0.00	12.00	2.00	12.00
	N	7.98	1.41	2.35	2.35	7.98		N	7.84	0.00	7.24	1.21	7.24
NS P5	R	14.00	8.00	1.00	1.00	2.00	NS P30	R	17.00	17.00	1.00	0.00	7.00
	N	11.69	6.68	0.83	0.83	1.67		N	8.15	8.15	0.48	0.00	3.36
NS P6	R	15.00	3.00	2.00	1.00	7.00	NS P31	R	35.00	2.00	1.00	13.00	12.00
	N	7.21	1.44	0.96	0.48	3.36		N	18.49	1.06	0.53	6.87	6.34
NS P7	R	27.00	1.00	9.00	3.00	9.00	NS P32	R	17.00	13.00	1.00	10.00	14.00
	N	12.36	0.46	4.12	1.37	4.12		N	8.26	6.31	0.49	4.86	6.80
NS P8	R	43.00	17.00	5.00	3.00	8.00	NS P33	R	40.00	3.00	4.00	5.00	14.00
	N	13.54	5.35	1.57	0.94	2.52		N	14.31	1.07	1.43	1.79	5.01
NS P9	R	31.00	8.00	17.00	24.00	7.00	NS P34	R	20.00	6.00	15.00	15.00	13.00
	N	8.27	2.14	4.54	6.41	1.87		N	7.70	2.31	5.77	5.77	5.00
NS P10	R	34.00	53.00	27.00	4.00	11.00	NS P35	R	19.00	14.00	4.00	9.00	4.00
	N	14.11	22.00	11.21	1.66	4.57		N	11.59	8.54	2.44	5.49	2.44
NS P11	R	27.00	3.00	38.00	8.00	11.00	NS P36	R	49.00	7.00	8.00	18.00	5.00
	N	10.56	1.17	14.87	3.13	4.30		N	26.11	3.73	4.26	9.59	2.66
NS P12	R	31.00	9.00	9.00	4.00	8.00	NS P37	R	17.00	10.00	19.00	4.00	5.00
	N	13.67	3.97	3.97	1.76	3.53		N	14.46	8.50	16.16	3.40	4.25
NS P13	R	32.00	17.00	0.00	2.00	18.00	NS P38	R	61.00	42.00	3.00	11.00	22.00
	N	13.90	7.38	0.00	0.87	7.82		N	20.79	14.31	1.02	3.75	7.50

NS P14	R	16.00	11.00	17.00	6.00	11.00	NS P39	R	34.00	9.00	12.00	14.00	15.00
	N	7.82	5.38	8.31	2.93	5.38		N	13.55	3.59	4.78	5.58	5.98
NS P15	R	44.00	4.00	26.00	15.00	0.00	NS P40	R	35.00	1.00	1.00	3.00	5.00
	N	12.88	1.17	7.61	4.39	0.00		N	20.05	0.57	0.57	1.72	2.86
NS P16	R	23.00	0.00	3.00	9.00	10.00	NS P41	R	13.00	28.00	16.00	1.00	13.00
	N	9.74	0.00	1.27	3.81	4.24		N	5.90	12.70	7.26	0.45	5.90
NS P17	R	14.00	6.00	10.00	3.00	7.00	NS P42	R	11.00	9.00	10.00	1.00	15.00
	N	7.08	3.03	5.06	1.52	3.54		N	5.52	4.52	5.02	0.50	7.53
NS P18	R	37.00	31.00	14.00	39.00	10.00	NS P43	R	20.00	3.00	42.00	30.00	15.00
	N	12.11	10.15	4.58	12.77	3.27		N	7.57	1.14	15.89	11.35	5.68
NS P19	R	31.00	16.00	26.00	24.00	14.00	NS P44	R	35.00	49.00	60.00	34.00	18.00
	N	11.36	5.87	9.53	8.80	5.13		N	13.20	18.48	22.62	12.82	6.79
NS P20	R	24.00	1.00	15.00	8.00	3.00	NS P45	R	33.00	6.00	12.00	7.00	6.00
	N	15.56	0.65	9.73	5.19	1.95		N	11.23	2.04	4.08	2.38	2.04
NS P21	R	10.00	4.00	4.00	0.00	15.00	NS P46	R	49.00	4.00	38.00	9.00	6.00
	N	4.44	1.78	1.78	0.00	6.66		N	15.60	1.27	12.10	2.87	1.91
NS P22	R	30.00	34.00	33.00	16.00	17.00	NS P47	R	31.00	1.00	1.00	3.00	1.00
	N	7.82	8.87	8.61	4.17	4.43		N	15.36	0.50	0.50	1.49	0.50
NS P23	R	16.00	1.00	4.00	0.00	8.00	NS P48	R	48.00	29.00	4.00	1.00	4.00
	N	11.21	0.70	2.80	0.00	5.61		N	20.56	12.42	1.71	0.43	1.71
NS P24	R	27.00	15.00	1.00	1.00	9.00	NS P49	R	29.00	21.00	6.00	1.00	16.00
	N	13.95	7.75	0.52	0.52	4.65		N	7.83	5.67	1.62	0.27	4.32
NS P25	R	34.00	17.00	13.00	7.00	11.00	NS P50	R	38.00	23.00	2.00	6.00	13.00
	N	14.93	7.46	5.71	3.07	4.83		N	14.88	9.01	0.78	2.35	5.09

The most frequent of the five discourse markers, *so*, was employed by all the participants. The participant (NS P36) with the highest frequency used it 26.11 times per 1,000 words. The frequencies of the remaining 49 participants ranged from 20.79 to 5.40 times per 1,000 words.

As for *like*, two participants (NS P16 and NS P29) never used it as a discourse marker, while NS P10 was the participant who employed it the most with a normalised frequency of 22.00 times per 1,000 words. The frequencies of the remaining 47 participants were between 18.48 and 0.46 times per 1,000 words.

With respect to *you know*, it did not occur in the speech of two participants (NS P13 and NS P27), whereas it was used by NS P44 the most with a frequency of 22.62 times per 1,000 words. The frequencies of the rest of the participants ranged between 16.16 and 0.50 times per 1,000 words.

Regarding *well*, it did not occur in the speech of only one participant (NS P15) during the interviews, whereas NS P2 employed it 11.97 times per 1,000 words, which is the highest frequency for this discourse marker. The frequencies of the remaining 48 participants varied from 7.98 to 0.50 times per 1,000 words.

Lastly, the least frequently occuring discourse marker in the LOCNEC, *I mean*, was never used by three participants (NS P1, NS P23 and NS P30) in their speech as a discourse marker, while NS P44 was the participant that employed it the most with a frequency of 12.82 times per 1,000 words. The frequencies of the remaining 46 participants were between 11.35 and 0.45 times per 1,000 words.

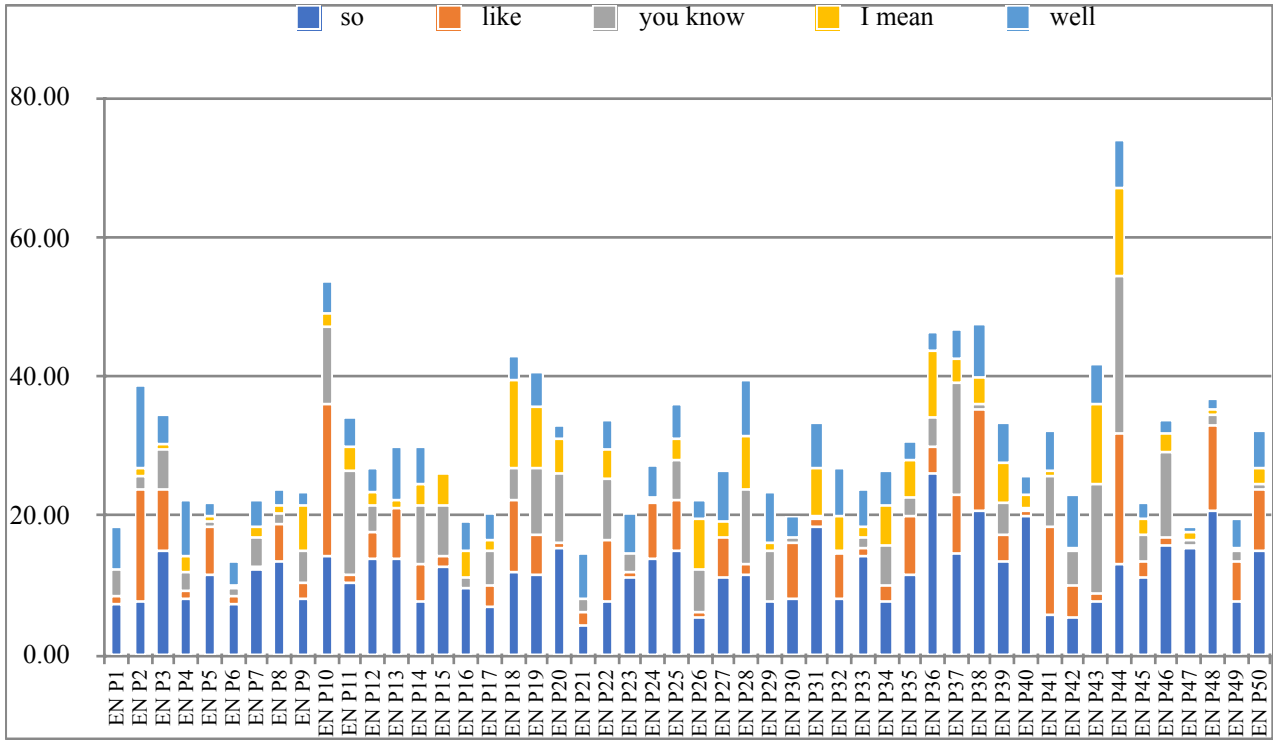


Figure 4.4. *Performance profiles of the British students for the five discourse markers (in per 1,000 words) in the LOCNEC*

As is seen in Figure 4.4, among the five discourse markers, *so* had the widest coverage within the LOCNEC in that all of the participants used it in their interviews. On the other hand, *I mean* was employed by the least number of participants; 47 out of 50 participants used it during the interviews, while three participants never used it. Moreover, *like* and *I mean* did not occur in the speech of only one participant each. To sum up, the distribution of the use of these discourse markers in the LOCNEC is far more balanced than is the case in the Turkish corpus.

4.4. Frequency Comparison of the Turkish Corpus and the LOCNEC

Based on the findings reported above, the use of the five discourse markers namely *so, well, you know, I mean* and *like*, by Turkish and native English speaking British students were compared in the Turkish corpus and the LOCNEC. In overall, with regard to the total normalised frequencies of this discourse markers, the British students (30.11 times per 1,000 words) employed these five discourse markers considerably more frequently than the Turkish students (18.46 times per 1,000 words) did. The comparison of the both groups in each of these discourse markers is represented in Figure 4.5.

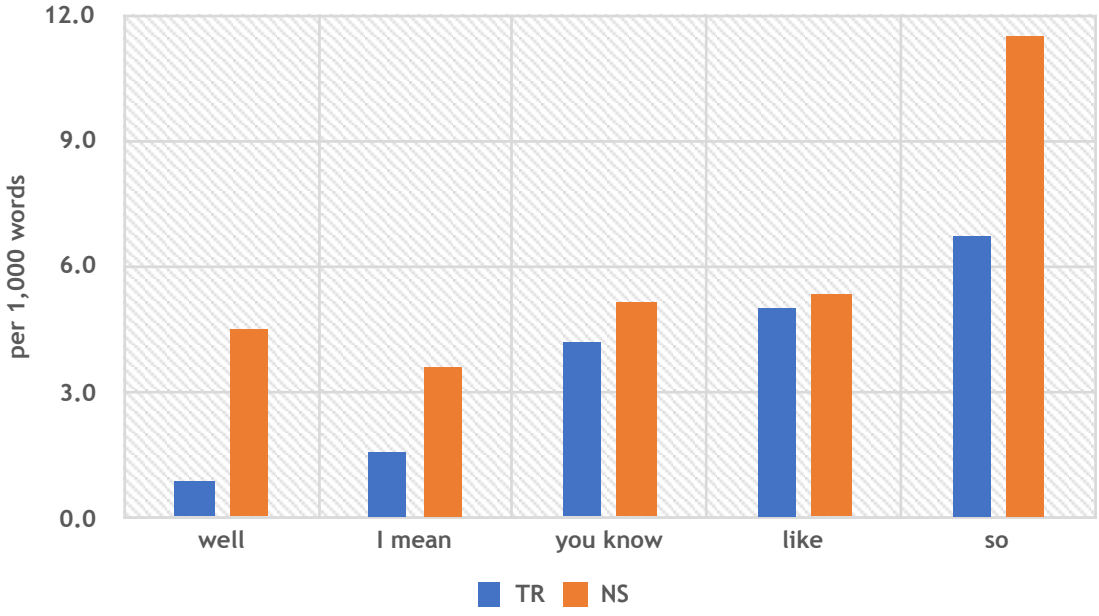


Figure 4.5. *Comparison of the discourse markers used by Turkish and British students*

The frequencies of these five discourse markers as used in the two corpora were also compared for statistical significance by means of the log-likelihood statistic and Mann-Whitney U test. The results are presented in Table 4.6 below.

Table 4.6. *Results of the log-likelihood statistic (G2) and Mann-Whitney U (U) test for the difference between the Turkish corpus and LOCNEC*

	So	*Like*	*You know*	*I mean*	*Well*
Log-likelihood	96.13	0.94	7.77	60.95	106.46
Sig. (*p*)*	<.0001	>.05	<.01	<.0001	<.0001
Mann-Whitney U	200.500	690.000	204.000	346.000	477.000
Sig. (*p*)**	<.0001	.0001	<.0001	<.0001	<.0001

* The log-likelihood wizard does not produce an exact *p* value, but gives the researcher critical G2 values for significance levels that are presented in the table.

** When the *p* values obtained in the Mann-Whitney U test went far beyond four decimals after the dot, so it was presented in the table as significance levels.

To begin with *so*, which is the discourse marker most frequently employed by both groups, the British students (11.52 times per 1,000 words) employed *so* considerably more frequently than the Turkish students (6.76 times per 1,000 words) did. According to the log-likelihood statistic and the Mann-Whitney U test, this difference was statistically significant ($G2$=96.13, p<.0001; U=200.500, p<.0001). In other words, the Turkish students significantly underused *so* as a discourse marker in their conversation compared to their native English speaking peers.

As for *like*, which is the second most frequently employed discourse marker in both groups, the British students (5.34 times per 1,000 words) used it slightly more frequently than the Turkish students (4.99 times per 1,000 words) did. Yet, this difference was not statistically significant ($G2$=0.94, p>0.05), which shows that the Turkish students as a whole group approximated to the use of their native peers in terms of frequency. However, the Mann-Whitney U test showed a statistically significant difference (U=690.000, p=.0001), which means that when the frequencies in the individual texts constituting both corpora, the difference in-between was significant. Apparently, some of the Turkish students used *like* quite frequently as a discourse

marker, while some never or rarely used it in such functions, which can be seen more clearly in the performance profile figures in the previous pages.

With regard to *you know*, the British students employed it 5.34 times per 1,000 words while the Turkish students used it slightly less frequently at 4.18 times per 1,000 words. Although the difference seems not to be large, it was statistically significant ($G2$=7.77, p<0.01, U=204.000, p<.0001), which means that the Turkish students significantly underused this discourse marker compared to their native peers.

Regarding *I mean*, the British students (3.62 times per 1,000 words) used it more than twice as much the Turkish students (1.60 times per 1,000 words) did in their interviews. This difference was also statistically significant according to the log-likelihood statistic and the Mann-Whitney U test ($G2$=60.95, p<0.0001; U=346.000, p<.0001). In other words, the Turkish students again significantly underused *I mean* in their conversation compared to their native peers.

Lastly, *well* as a discourse marker occurred in the speech of the British students (4.48 times per 1,000 words) almost five times more than in that of the Turkish students (0.90 times per 1,000 words). This particularly large difference was statistically significant as well ($G2$=106.46, p<0.0001; U=477.000, p<.0001), which again shows that the Turkish students significantly underused this discourse marker compared to their native peers.

4.5. Comparison of First- and Fourth-Year Students in the Turkish Corpus

In the scope of this study, the first- and fourth-year students interviewed for the Turkish corpus were also compared in terms of their use of the five discourse markers, namely *so, well, you know, I mean* and *like*. The raw and normalised frequencies of the Turkish participants in both groups are presented in Table 4.7 below.

Both first- (8.20 times per 1,000 words) and fourth-year (5.38 times per 1,000 words) students employed *so* as a discourse markers with the highest frequency in their interviews, while using *well* with the lowest frequency (1.09 and 0.73 times per 1,000 words, respectively).

Table 4.7. *Frequencies of the discourse markers in the speech of the first- and fourth-year Turkish students*

		so	like	you know	I mean	well
	IS: 29,376 words					
First	Raw	241	203	148	59	32
	Normalised	8.20	6.91	5.04	2.01	1.09
	IS: 30,311 words					
Fourth	Raw	163	95	102	37	22
	Normalised	5.38	3.13	3.37	1.22	0.73

IS: Interviewee Speech

In overall, the first-year Turkish students (23.25 times per 1,000 words) employed these five discourse markers more frequently than the fourth-year students (13.83) did. The comparison of the both groups in each of these discourse markers is represented in Figure 4.6.

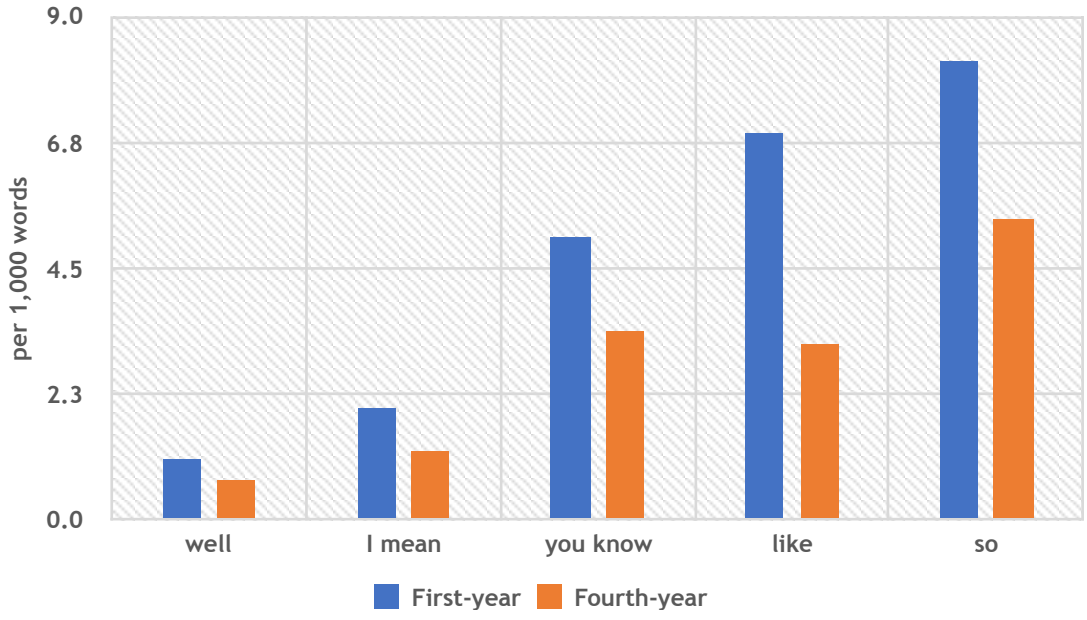


Figure 4.6. *Comparison of the discourse markers used by first- and fourth-year Turkish students*

Although both first- (8.20 times per 1,000 words) and fourth-year (5.38 times per 1,000 words) students employed *so* as a discourse markers with the highest frequency in their interviews, the first-year students' use of discourse markers was considerably more

frequent than those of the fourth-year students. The differences in-between were compared by means of the log-likelihood statistic and Mann-Whitney U test, and the results are presented in Table 4.8.

Table 4.8. *Results of the log-likelihood statistic (G2) and Mann-Whitney U (U) test for the difference between the first- and fourth-year Turkish students*

	So	*Like*	*You know*	*I mean*	*Well*
Log-likelihood	17.70	43.50	10.02	5.80	2.19
Sig. (*p*)*	<.0001	<.0001	<.01	<.05	>.05
Mann-Whitney U	231.000	303.500	267.000	233.500	298.500
Sig. (*p*)	.113	.860	.238	.053	.780

* The log-likelihood wizard does not produce an exact *p* value, but gives the researcher critical G2 values for significance levels that are presented in the table.

According to log-likelihood statistic, this difference was statistically significant (*G2*=17.70, *p*<0.0001), but the Mann-Whitney U test did not reveal a significant difference (*U*=231.000, *p*=.113). Such a difference was also observed in the use of *like* in that the first-year students employed it 6.91 times per 1,000 words while the fourth-year students used it 3.13 times per 1,000 words. This difference was also statistically significant in the log-likelihood statistic (*G2*=43.50, *p*<0.0001), but not significant in the Mann-Whitney U test (*U*=303.500, *p*=.860).

The difference between the first- and fourth-year students in the use of *you know* was slightly smaller compared to *so* and *like*. *You know* occurred 5.04 times per 1,000 words in the speech of the first-year students, and 3.37 times per 1,000 words in that of the fourth-year students. However, the difference was still statistically significant at the level of *p*<0.01 (*G2*=10.02) in the log-likelihood statistic, although it was not significant in the Mann-Whitney U test (*U*=267.000, *p*=.238).

The smallest differences between the two groups of Turkish students were observed in the use of *I mean* and *well*. The first-year students used *I mean* 2.01 times per 1,000 words while the fourth-year students used it 1.22 times per 1,000 words. As for the log-likelihood result, this difference was revealed to be statistically significant at the level of *p*<0.05 (*G2*=5.80), and the Mann-Whitney U test did not show a significant

difference (*U*=233, *p*=.053). On the other hand, *well* as a discourse marker occurred 1.09 times per 1,000 words in the speech of the first-year students, and 0.73 times per 1,000 words in that of the fourth-year students. The difference in the frequency of *well* between the two groups was not statistically significant according to both the log-likelihood statistic and Mann-Whitney U test (*G2*=2.19, *p*>.05; *U*=298.500, *p*=.780).

In both statistical tests, *well* was the only discourse marker that did not reveal a significant difference, but in other discourse markers, the log-likelihood statistic showed a significant difference while the Mann-Whitney U test did not. This is probably due to the fact that the data varied too much when individually evaluated, and in that case there was no significant difference between the first- and fourth-year students (as revealed by the Mann-Whitney U test). On the other hand, when evaluated as a whole (or in total), the two groups differed significantly in four of the five discourse markers in terms of frequency. Above all, the first-year students made use of the five discourse markers considerably more frequently than their fourth-year peers.

4.6. Functions of Discourse Markers in the Turkish corpus and the LOCNEC

The functional analysis of the discourse markers occuring in the Turkish corpus and the LOCNEC is presented in this section. The findings are then elaborated for each function and sample quotations from the Turkish corpus and LOCNEC are presented accordingly.

4.6.1. So

The Turkish students who were interviewed for the Turkish corpus and their native English speaking peers from the LOCNEC employed *so* for a variety of functions, with certain degrees of differences in the proportions of these functions in the two corpora. The overall findings are presented in Table 4.9 below.

In terms of overall frequency, there was a considerably large difference in the frequencies of *so* between the two corpora (404 times in the Turkish Corpus vs. 1364 times in the LOCNEC).

Table 4.9. *Functions of so in the Turkish Corpus and LOCNEC*

	Turkish Corpus		LOCNEC		Significance (X2/G2)
	n	%	n	%	
Textual					
Indicate a result	144	35.64	383	28.08	$p<.01$
Introduce a summary	64	15.84	200	14.66	n.s.
Move back to a previous discourse unit	24	5.94	49	3.59	n.s.
Introduce the next sequence	32	7.92	84	6.16	n.s.
Provide elaboration	16	3.96	90	6.60	n.s.
Mark a discourse boundary	0	0.00	7	0.51	n.s.
Subtotal	280	69.31	813	59.60	$p<.001$
Interpersonal					
Draw a conclusion	64	15.84	145	10.63	$p<.01$
Mark an opinion	12	2.97	71	5.21	n.s.
Subtotal	76	18.81	216	15.84	n.s.
Interactional					
Hold the floor	24	5.94	90	6.60	n.s.
Give the floor	16	3.96	220	16.13	$p<.001$
Subtotal	40	9.90	310	22.73	$p<.001$
Unclassified Instances	8	1.98	25	1.83	
Total	404	100	1364	100	

As for the functions these instances served in discourse, a total of 10 functions were identified in the corpora examined in this study. In both corpora, most of the instances (about 69% in the Turkish corpus and about 60% in the LOCNEC) functioned in the textual domain. It was followed by the interpersonal and interactional domains, respectively, for the Turkish corpus, but the instances that functioned in the interactional domain outnumbered those that functioned in the interpersonal domain for the LOCNEC. This means that the British students employed *so* in interactional functions more frequently than the Turkish students did, whereas the Turkish students used this discourse marker in textual functions more frequently than their British peers.

When the differences in individual functions are examined, the findings reveal significant differences, as is seen in Table 4.9, between the Turkish and British students when using *so* to indicate a result ($p<.01$), draw a conclusion ($p<.01$), and give the floor ($p<.001$). The Turkish students employed *so* to indicate a result and draw a conclusion significantly more frequently than the British students, whereas the British students used it to give the floor to the interlocutor significantly more frequently than the Turkish students. What all the functions reported in Table 4.9 mean and do in discourse are presented in the following sub-sections along with sample quotations from each corpus for every function revealed.

4.6.1.1. So to indicate a result

To indicate a result or consequence is probably the most common function that is attributed to *so* in the literature (Müller, 2005; Buysse, 2012; Blakemore 1988; Schiffrin 1987; etc.), which is also reflected in the raw frequencies of the instances that served this function in the study, i.e. the most frequently employed function in both corpora. According to Biber et al. (1999:877), *so* has a typical resultative meaning, while Schiffrin (1987) argues that other functions of *so* are an extension of this core "resultative" meaning. This function basically means that what is said is that result or consequence of what is said before; Buysse (2012: 1768) puts it this way: "state of affairs Y is the result/consequence of state of affairs X". Examples of this function from the corpora examined are provided below.

(9) <B> but erm I was gonna go to Hull but I didn't get a good grade in one of my A-levels . *so* they turned me down and I thought well I'm just gonna reapply . and then after I'd worked for a while I thought oh I'd like to <X> management <\B> NS Participant 35

In Example (9), the British interviewee mentions a school at which she was planning to study, but apparently did not meet the admission criteria and that is why she says she was turned down. Here, *so* indicates a result; being turned down due to not having a good grade.

(10) <B> […] and I waited there hungry and thirsty until the morning Vienna is too too expensive I can't afford living there I realised it . (er) until morning my bus came at at seven o'clock so I . I (erm) was in Prague after five hours *so* I didn't go to the work . my boss really mad at me and she said why didn't you tell us (er) </B> TR Participant 27

Likewise, in Example (10), the Turkish interviewee is talking about a journey to Vienna during her Erasmus internship experience. Since her bus was cancelled, she had to wait at the bus station for hours. Through the end of the turn, she says that she could be at work on time, which is marked by *so* at the beginning of the utterance. This is because her not being able to be at work on time is a result of the bus coming very late.

4.6.1.2. So to introduce a summary

In this function of *so*, it presents a segment that sums up the prior discourse (Redeker, 1990; Müller, 2005), and does not provide a new claim regarding the previous text, but simply restates the main argument that can be inferred from the prior discourse in more general terms (Buysse, 2012). In this function, the utterance that follows *so* indicates the same message as a previous utterance produced early in the same turn, or a few turns back. Examples of this function from the corpora examined in the present study are as in the following:

(11) <B> er but Warrington is a lot richer in terms of work and .. things like that I mean it's a richer full stop I mean <X> very much er <\B>
<A> you can't get bored <laughs> <\A>
<B> no I mean there's a big social life <\B>
<A> [mhm <\A>
<B> [I mean there's about nine night-clubs <\B>
<A> oh yes [mhm <\A>
<B> [yeah and about I don't know about twenty pubs in the town centre <\B>
<A> mhm <\A>
<B> *so* there's a big social life going on erm .. and because <X> so many motor ways lots of people have built their head offices there so <X> British Nuclear <XX> <\B>
NS Participant 36

In Example (11), the British interviewee is telling how life is like in a city called Warrington that is a few hours south of Lancaster where the interviews for the

LOCNEC were conducted. After describing the city as having a big social life, she provides information regarding the number of night clubs and pubs located in the town centre. Then, she sums up the issue by stating the main point, i.e. *big social life in Warrington*, which is marked by *so* and moves on to another aspect that makes the city a central place.

(12) <B> […] and also I was really impressed (er) by the church there was big church and . (erm) there were a ceremony you know like Christians <X> it actually scared me because it was so like haunting and you know like there was a lot of people they were serious and you know the voice were echoing because it was so big *so* it kind of scared me a little bit but . it was good because architecture and you know like the walls and the pictures were really impressive […] <\B> TR Participant 16

Similarly, in Example (12), the Turkish interviewee describes her visit to a church in a historical town as scary, and then explains why it looked that scary. Afterwards, she comes back to the main point that the atmosphere scared her, which is marked by *so*, and then moves on to a positive aspect of this visit, i.e. the architecture being good as well as the walls and pictures being impressive.

4.6.1.3. So to move back to a previous discourse unit

In this function, *so* can indicate a shift to a higher textual level, which can be either after a brief interruption due to an arising issue/dimension or an exchange with the interlocutor (Buysse, 2012). Here, *so* can mark a resumption (Polanyi and Scha, 1983, cited in Buysse, 2012, p. 1773), i.e. beginning something after a pause or interruption, mark a main idea unit (Müller, 2005; Schiffrin, 1987), indicate a return to the main story (Vandepitte, 1993), and mark a return from a parenthetical segment (Redeker, 2006). Examples of this function from the corpora examined in the present study are as in the following:

(13) <B> you have to you don't have a choice so I had to find my way around the town I had to use [<XX> <\B>
<A> [where was it <\A>
<B> it was <name of town> it was down in the south west of France <\B>

<A> mm [nice and warm <\A>
<B> [erm . nice and warm yes near Albi you know Albi <\B>
<A> mhm <\A>
<B> erm .. *so* I had to go shopping I had to go to the supermarket I'd never been to the supermarket on my own I'd always been with friends when I was here cos it's […] </B>
NS Participant 34

In Example (13), the British student is telling her experience of being alone in a different city for the first time. She first says that she has no choice, but to find her way around the town. However, then, the interviewer wonders where that city is, and there is some back-and-forth communication about it and how it is like. Afterwards, the interviewee wants to go back to where she left before the digression, and employs *so* and continues to tell her experience.

(14) <A> do we have a church I didn't know <laughing> </A>
<B> (er) I have a Christian friend at the dormitory . I had a Christian friend and he took me to the church there was an </B>
<A> where is the church </A>
<B> I don't know where is (er) Ataturk high school you know </B> <A> ha yeah in that region in Odunpazari </A>
<B> in that area . *so* there was an American girl I talked to her (er) he want she wanted to meet me but I didn't want because I <laughing> how can I say (erm) . my friend and I are not friends any more […] </B> TR Participant 13

Comparably, in Example (14), the interviewer previously asked the Turkish interviewee how she improved her English, and she said she tried to chat with native speakers as much as possible. Here, she is telling how she met a native English speaker in Eskisehir. Apparently, she met her in a church, but at that moment, the interviewer pops in because she did not know there was a church in the city centre. After informing him about the whereabouts of this church, she goes back where she left by employing *so* and continues to tell how they met.

4.6.1.4. So to introduce the next sequence

This function basically helps achieve a coherent transition from one event to another within a narrative, or introduce the next part of a story. According to Redeker

(1990), *so* has potential of marking sequential relations ‘‘between successive elements in a chain of events’’ (1990:373–374). In other words, the sequential *so* starts a new sequence within a turn rather than bringing a sequence or a turn to a closure (Buysse, 2012). Examples of this function from the corpora examined in the present study are as in the following:

(15) <B> […] she goes to: . to a friend who she knows and asks him to paint the portrait . which he does . and he paints it exactly as he sees it . and he says right I've <?> done come and have a look at this . *so* she gets up and she looks and she's absolutely flabbergasted and she says .. but but but I don't look like that . cos he hasn't painted her in a particularly nice light […] </B> NS Participant 46

In Example (15), the British interviewee is narrating the story that she made up based on the four pictures shown at the end of the interview. Sequences within this story can be seen in the turn provided above. When the interviewee employs *so* in the middle of the turn, she then provides a new sequences, i.e. *she gets up and she looks* which is apparently sequenced after the painter asked her to *come and have a look*.

(16) <B> yeah I .. once upon a time there was a selfish . concited woman who . want to . </B>
<A> it doesn't have to be like a past time story </A>
<B> wants to . want a picture of herself *so*: she goes to a painter and she asks him to paint herself . and </B>
<A> like a portrait </A>
<B> yes like a portrait *so*: . he says okay I'll do that <laughing> and he draws her . but she doesn't like it *so*: she wants to take the picture to her friends . and she asks them if it's . if it represents the reality . like am I really like this am I this ugly . no I'm the . most perfect girl in the world <laughing> […] </B> TR Participant 1

Likewise, in Example (16), the Turkish interviewee is also talking about the picture-description task at the end of the interview, which makes sense because sequences usually emerge in stories. She uses *so* multiple times throughout two different turns to introduce the next sequence in the story she is narrating.

4.6.1.5. So to provide elaboration

This function is about providing further information regarding an utterance that has just been produced. It occurs when ‘‘one clause elaborates on the meaning of another by further specifying or describing it’’ (Halliday and Matthiessen, 2004:396). In this function, *so* precedes an explanation regarding a prior more general statement, and this explanation is usually non-argumentative, but merely supplements information provided only just, and it can also be inserted parenthetically in a turn (Buysse, 2012). Examples of this function from the corpora examined in the present study are as in the following:

(17) <B> because once you've got children you always have to plan in for the[i:] unexpected because <\B>
<A> yes <\A>
<B> if they're ill or something <\B>
<A> <X> <\A>
<B> then they want [you all the time <\B>
<A> [you've got to stay there mhm <\A>
<B> and you you've got to give them that time *so* you've always got to have a little bit of of leeway you can't be doing two essays in two days <\B>
<A> mhm <\A> NS Participant 45

In Example (17), the British interviewee is talking about the downsides of having children while doing a Master’s degree. She argues that because you have children, you have to plan for the unexpected all the time, and when they want you, you have to give them that time. However, this does not directly relate to doing an Master’s with children, and she employs *so* and then, elaborates further by rewording her main argument and providing an example.

(18) <A> exactly (er) so your experience in Poznan how was it </A>
<B> (er) this was not good as (er) Spain because (er) I went to Poland (er) during my second year here in university *so* (er) after one year studying here my English was not good at was not good to communicate with the people I had an Italian friend and he helped me a lot to practise my English and this was basically going to classes sometimes not going because […] </B> TR Participant 35

Similarly, in Example (18), the Turkish interviewee is talking about her Erasmus experience in Poland, and whether it was good or not. She initially says that it was not good, and explains why. The reason she provides is that it was her second year at university, but this clearly needs to be elaborated because the hearer needs to understand why being a second year university student would make an Erasmus experience not good. Then, she employs *so* and provides that elaboration that seems to be needed.

4.6.1.6. So to mark a discourse boundary

In this function, *so* can indicate the opening of a new section in the conversation, and it is reported in studies that especially have data that are similar to those in this study, i.e. *interviews*, or involving discussions with a set of tasks (e.g. Müller, 2005; Buysse, 2012). In the present study, it occurred particularly at the beginning of the interviews where the interviewees had some time of thinking and then started talking about the topic they chose, or at the end of the interviews where they were asked to describe the pictures provided to them. The instances that served this function of *so* were mostly in these parts of the interviews. Examples of this function from the corpora examined in the present study are as in the following:

(19) <A> so which topic have you chosen <\A>
<B> topic number two <\B>
<A> number two so it's a country <\A>
<B> yeah <\B>
<A> yes <\A>
<B> *so* em . I went to Cuba in nineteen ninety-two <\B>
<A> right <laughs> <\A> NS Participant 47

In Example (19), the British interviewee has looked at the topics, and after thinking about them, she chooses to talk about a country that she visited and that impressed her. After stating which topic she chose, she starts speaking about it, which is a new section in the conversation, and is marked by *so* at the beginning. Similar examples were found at the beginning of the picture-description tasks at the end of the

interviews where the interviewees started talking about the story that they made up after sorting it out.

4.6.1.7. So to draw a conclusion

This function denotes a relation of conclusion between two utterances; the one that comes the first functions as the base for a claim in the second. It was presented in Table 4.7 as functioning at the interpersonal domain because it involves some personal involvement form the speaker in the message conveyed. Examples of this function from the corpora examined in the present study are as in the following:

(20) <A> so you're brave <laughs> <\A>
<B> <laughs> it's okay coming in cos it's all down hill but er coming back home it takes forever <\B>
<A> .. how long does it take <\A>
<B> it takes about ten minutes coming in and more than fifteen going back it's up near the[i:] [Ashton Memorial<\B>
<A> [well .. oh oh yes <\A>
<B> *so* [that's quite steep <laughs> <\B>
<A> [oh yeah that is quite <\A> NS Participant 40

In Example (20), the British interviewee is talking about cycling to a place, which is apparently up hill. After saying that it takes longer to go back from that place, he draws the conclusion that the route is quite steep because it takes much more time.

(21) <B> yeah Indian movie <laughing> it was long before I watched it I . think that I won't gonna like it because Indian movies like three hours . and then I began to watch it and was really really impressed you know I saw that it was about discrimination of the Muslims and I was really impressed because I had no idea . of that something like that exist before I watched it </B>
<A> really </A>
<B> yeah <laughing> I didn't know but then I realised that . it's true and I see I began to see that (er) . I began to experience it . for example I .. (erm) I looked at an online website to find some international friends . and <laughing> most of . most of them blocked Turkey <laughing> they don't answer my messages and . I don't know I have native friends here they are really good but <X> (er) some of them are didn't know don't know us . *so* they are

really prejudiced about us (er) I talked someone from Netherlands you know . lately our relationship is not really well <laughing> </B> TR Participant 13

In like manner, in Example (21), the Turkish interviewee states that she realised there was discrimination against Muslims when she watched a movie. As the interviewer seems surprised, the interviewee elaborates on this observation from her experience. Based on this experience, she at the end draws that conclusion that it is prejudice which is marked by *so*.

4.6.1.8. So to mark an opinion

In this function, *so* precedes an opinion stated by the speaker, in that he/she expresses what he/she thinks by usually using an adjective. This opinion is presented based on what is said just before, and also includes an element of result as well (Buysse, 2012). Examples of this function from the corpora examined in the present study are as in the following:

(22) <B> and how every one was trying to: to take money if they could and the power relationship between the[i:] aristocracy <\B>
<A> oh yes <\A>
<B> and who they were supporting whether they were supporting the[i:] English or whether they were supporting the Scots erm even if they were Scottish it was quite accurate *so* that was a good film but I just didn't think that it it it carried so much weight […] <\B> NS Participant 43

In Example (22), the British interviewee is talking about a movie that impressed him, and after describing what it was about, she then states his overall opinion about that movie, i.e. *that was a good film*.

(23) <A> so (er) the courses were not really difficult for you then </A>
<B> no no there was one there was this one course and they taught us if clauses and I said you know I teach this to my students and they are now teaching this to me *so* it was kind of boring the courses were really boring </B> TR Participant 36

In a similar vein, in Example (23), the Turkish interviewee compares the difficulty of courses in her home university and the university she studied at during Erasmus programme. She thinks that the courses were easier at the university she visited, and also gives an example from a course. At the end of the turn, she employs *so* and states her opinion of this experience.

4.6.1.9. So to hold the floor

In this function, *so* precedes and/or follows filled or unfilled pauses, and pronounced with a prolonged vowel (i.e. represented with a colon in transcriptions). According to Lam (2009), this is a "processing" function, and can be used "as a delay strategy and signal that the speaker is undergoing some processing problem and requires extra time" (2009:364). Therefore, the speaker has an intention to continue speaking, but needs some time for processing and figuring out what to say or the organisation of the following utterance. Examples of this function from the corpora examined in the present study are as in the following:

(24) <B> so he had to learn Greek <\B>
<A> yes but modern Greek <\A>
<B> modern Greek cos Ancient cos Greeks don't speak Ancient Greek <\B>
<A> [mhm <\A>
<B> [just like Italians don't speak Latin .. erm *so:* . he went he's learnt to speak Greek and then he did a course teaching English as a foreign language […] <\B>
NS Participant 21

In Example (24), the British interviewee previously mentions a friend who wants to go to Greece to teach English. At some point within a turn, he gives a brief unfilled and then a filled pause, which is followed by *so* with a prolonged vowel and then another brief pause. In this example as well, the speaker seems to buy some time by using *so* and at the same time hold the floor so that he could continue speaking after figuring out what to say.

(25) <B> also there are lots of motives Turkish motives (er) on its (er) . basilica this basilica . so it's really important and really interesting for me .. and . I don't remember anything . *so:* . there are lots of birds in that square <laughing> I remember that . and </B>
TR Participant 3

Likewise, in Example (25), the Turkish interviewee is talking about a city that she visited and how historical it was. At some point, she tries to go on speaking and provide more information about that city, but apparently she has difficulty in retrieving that information or uttering what she has in mind. Still, she intends to continue and employs *so* with a prolonged vowel, which is preceded with and followed by a brief unfilled pause.

4.6.1.10. So to give the floor

Unlike the previous function, i.e. to hold the floor, this function indicates an implied result and a desire to give the floor to the interactant. Thus, *so* in this function occurs at the end of a turn and is usually followed by an unfilled pause and gives the hearer the clue for taking over the floor (Müller, 2005). In other words, it marks a potential turn-transition, and expresses an individual's desire to take a more passive role so that he/she could give the floor to the addressee (Lam, 2010). Examples of this function from the corpora examined in the present study are as in the following:

(26) <A> [but I didn't know that it was based on a true story as well <\A>
<B> [but the ... yes Ken= Kensington <X> no that's Mi= Miss Kensington is the house keeper can't remember the name of the place <\B>
<A> mm I couldn't tell you I'm not very good at names especially in films <\A>
<B> well it's one of my favourite films *so* <\B>
<A> mhm do you often go to the cinema <\A>
<B> not so much these days I used to a lot . erm well over the past year or so I haven't […] <\B> NS Participant 21

In Example (26), the British interviewee is talking about a movie, and tries to remember the name of a place although it is one of his favourite movies. Since he cannot remember it, he apparently does not have anything else to say about this place, and ends the turn with *so* presumably to give the floor to the interviewer.

(27) <A> so you go there every year </A>
<B> every year we try to go but . now we can't actually because I have this university thing going on and summers are busy for me because I mostly work *so* </B>
<A> so where do you work in summers </A>
<B> I'm actually a computer engineer not licensed but I know programming to some degree </B> TR Participant 5

In parallel, the Turkish interviewee in Example (27) answers a question, i.e. *whether he goes to a place every year*. After he provides sufficient information in this answer, it seems that he uses *so* at the end of the turn to indicate that that is all and hand over the floor to the interviewer.

4.6.2. Like

The Turkish students who were interviewed for the Turkish corpus and their native English speaking peers from the LOCNEC employed *like* for a variety of functions, with certain degrees of differences in the proportions of these functions in the two corpora. The overall findings are presented in Table 4.10 below.

Table 4.10. *Functions of like in the Turkish Corpus and LOCNEC*

	Turkish Corpus		LOCNEC		Significance (X2/G2)
	n	%	n	%	
Textual					
Approximate number/quantity	67	22.48	125	19.75	n.s.
Introduce an explanation	55	18.46	83	13.11	$p<.05$
Introduce an example	46	15.44	142	22.43	$p<.05$
Quotative	55	18.46	75	11.85	$p<.01$
Mark lexical focus	50	16.78	142	22.43	n.s.
Subtotal	273	91.61	567	89.57	*n.s.*
Interactional					
Search for the appropriate expression	21	7.05	33	5.21	n.s.
Restart/False start	0	0	8	1.26	*n.s.*
Subtotal	21	7.05	41	6.48	*n.s.*

Unclassified Instances	4	1.34	25	3.95
Total	298	100	633	100

In terms of overall frequency, there was a considerably large difference in the frequencies of *like* between the two corpora (298 times in the Turkish Corpus vs. 633 times in the LOCNEC). As for the functions these instances served in discourse, a total of seven functions were identified in the corpora examined in this study. In the both corpora, most of the instances (about 92% in the Turkish corpus and about 90% in the LOCNEC) functioned in the textual domain.

When the differences in individual functions are examined, the findings reveal significant differences, as is seen in Table 4.10, between the Turkish and British students when using *like* to introduce an explanation ($p<.05$), introduce an example ($p<.05$), and provide a quotation ($p<.01$). The Turkish students employed *like* to introduce an explanation and provide a quotation significantly more frequently than the British students, whereas the British students used it to introduce an example significantly more frequently than the Turkish students. What all the functions reported in Table 4.10 mean and do in discourse are presented in the following sub-sections along with sample quotations from each corpus for every function revealed.

4.6.2.1. Like to approximate number/quantity

As noted by many researchers (e.g. Andersen 1997, 1998; Schourup 1985), a primary function of *like* is to indicate approximation or looseness of meaning. It frequently occurred in the data and findings of Andersen (2001) and Schourup (1985) as well as many others. In this function, *like* preceding a numeral expression, it indicates that the number or quantity is not an exact one, and thus, "reduces the speaker's commitment to the literal truth of his/her utterance" (Müller, 2005: 210). Examples of this function from the corpora examined in the present study are as in the following:

(28) <A> I suppose you've been teaching <\A>
<B> yeah <\B>
<A> for [a number of years and <\A>

<B> [erm .. well . not that long I've only got *like* four years of full-time teaching experience <\B> NS Participant 33

In Example (28), the British interviewee uses *like* to state an approximate, not an exact, number as she is talking about the amount time during which she had been teaching.

(29) <B> yeah but <laughing> but I will of course say United States because you know you know that we have been learning English for *like* twelve years and . and many of us can't find an opportunity to go abroad and […] </B> TR Participant 31

Similarly, in Example (29), the Turkish interviewee is talking about a country that impressed him, and why it was particularly impressive. He refers to the fact that they had been learning English for a long time, but did not have an opportunity to visit an English speaking country. At this point, as he refers to the amount of time (i.e. years) during which he had been learning English, he employs *like* just before the number to indicate that it is approximate, not exact.

4.6.2.2. Like to introduce an explanation

In this function, the speaker uses *like* because he/she wants to extend the information he/she has just provided to make it more understandable, or in other words, gives an explanation of that information (Müller, 2005). This can be in the form of providing additional information or introducing a further elaboration. Examples from the corpora examined in the present study are as in the following:

(30) <B> and they did it in in style . and you have had to have a good time <\B>
<A> mhm . what do you mean they did it in style <\A>
<B> well you know <X> sort of walking through . Rio <\B>
<A> mhm <\A>
<B> and erm .. it's just they just closed off the street and had a party in it *like* had a beauty contest bands going and […] </B> NS Participant 41

In Example (30), the British interviewee is talking about a past experience and describing the scene that she witnessed on a street. She mentions *a party* that is apparently going on, but then uses *like* and provides an explanation of the information that was just communicated.

(31). <A> so can you speak French </A>
<B> no I can't <laughing> </B>
<A> but you were taught </A>
<B> yeah they taught me . they also offer French classes as well *like* normal school curriculum but the teachers teach in French if you choose you can go to those kind of schools I didn't .. what else </B> TR Participant 41

Likewise, in Example (31), the Turkish interviewee employs *like* just after he communicates new information and before he provide an explanation regarding it. He is talking about his French learning experience as he lived in Canada for a certain amount of time. He says students are offered French classes at school there, and provide additional information related to it by stating that it is within the normal school curriculum.

4.6.2.3. Like to introduce an example

Another function of *like* that has been frequently reported in the literature is to introduce an example that seems to "have the meaning of 'for example'" (Schourup, 1985: 48). This example can be of a general concept that is not usually stated in the immediate context, or of an incident experienced before. Examples of this function from the corpora examined in the present study are as in the following:

(32) <B> […] you know it's surrounded by the fields and stuff but I suppose there's not really much way around it unless you put *like* a lake in the middle of the campus or something like that <\B> NS Participant 44

In Example (32), the British interviewee does not think the university campus is really attractive because of concrete buildings. She describes how it could be attractive

with an example, i.e. *putting a lake in the middle of it*, and this example is marked by *like*.

(33) <B> […] because I went there with my class with the school I had my friends and we wanted to go *like* sea places go to shop=go shopping and stuff and they were like we don't wanna we don't feel like going so we don't wanna go and we depended on them if they didn't wanna go we couldn't go . so it was . yeah that was very mean <laughing> </B>
TR Participant 7

Similarly, in Example (33), the Turkish student is talking about a school trip during her high school education. She thinks that they were not free during the trip because they wanted to go out, and provides examples of where they would want to go. At this point, she uses *like* to provide an example.

4.6.2.4. Like to introduce a quotation

In this function, *like* marks the beginning of a quotation; this can be from a hypothetical conversation or self-talk, or a real conversation that the speaker has had before. It usually, but not necessarily, forms the part of a construction *be+like*, and is integrated into the syntactic structure of a sentence. However, there seems to be some disagreement among scholars regarding the discourse-marker status of *like*. Some scholars including Fuller (2003), Dailey-O'Cain (2000) and Müller (2005) deny discourse-marker status to this usage since it cannot be omitted without making the utterance ungrammatical, whereas Jucker and Smith (1998: 183) and Andersen (1998: 156) include this function of *like* as one of its discourse marker functions. In this study, it is also included as a discourse-marker function because it is clearly a different usage considering the non-discourse-marker usages of *like*, and the quotation use of *like* can also occur alone, i.e. not part of the *be+like* construction. Examples of this function from the corpora examined in the present study are as in the following:

(34) <B> […] in front of it yeah this was different <?> made you a bit ashamed to be English I was *like* oh no no no I'm Scottish really <X> <\B>
<A> <laughs> <XX> in Scotland <\A> NS Participant 49

In Example (34), the British interviewee was previously talking about a movie in which an English character is attacking others, presumably Scottish, and the interviewee provides a quotation from a conversation that might have happened while watching that movie with Scottish friends.

(35) <B> […] people even if they didn't know you . you are going straight they coming from the other way they see they are *like* oh how are you hello how is your day this is the weather is really nice they make a conversation with you very friendly and the neighbourhood I lived in was very safe […] <\B> TR Participant 7

Likewise, in Example (35), the Turkish interviewee is talking about her experience in the US, and how people were like there. She provides a quotation from a hypothetical conversation with an American in ordinary, everyday life. This quotation is marked by *like* just before it.

4.6.2.5. Like to mark lexical focus

Another function, *like* as a focuser, marks new information or focus (Buysse, 2012), and Underhill (1988) argues that 'focus' is "the most significant new information in a sentence - often the point of the sentence (p. 234), and sets off unusual notions. According to Müller (2005), *like* can focus on new information as well as given information. On the other hand, Andersen (2001) refers to this function as meta-linguistic focus, and states that "like can be construed as a signal that the chosen expression does not fit readily into the linguistic repertoire of the speaker, i.e. that the speaker feels a minor discomfort with its use" (2001: 243). Examples of this function from the corpora examined in the present study are as in the following:

(36) <B> so you have to go and swim <\B>
<A> and do you wear wet suits or <\A>
<B> oh no a swimming costume . in October usually it's usually in October . we did ours in October it was cold it was so cold that halfway round you can see why people drown cos you just want to stop you just think I don't wanna swim any more people come out with *like* bruises all over their legs where they've hit rocks at the bottom and just didn't know . they just didn't feel it <\B> NS Participant 24

In Example (36), the British interviewee is talking about sort of a religious ceremony in which people swim in cold weather. One can barely assume that people would normally have bruises on their legs after swimming; it is something unusual. Moreover, this use of *like* here in this example does not correspond to *approximately* or *for instance* because there are no other signs of injury in the on-going discourse. Rather, it seems that the information presented after *like*, i.. *bruises*, is the focus of the utterance, and not an approximation.

(37) <A> how long did you stay there </A>
<B> three days . three or four days I think </B>
<A> okay . it's not too bad <laughing> </A>
<B> we also found a Turkish restaurant <laughing> it is also very good for me <laughing> because I was in Erasmus and I was *like* hungry for twenty years . because in Lithuania the cuisine are very different from Turkey </B> TR Participant 34

Similarly, in Example (37) above, the Turkish student is talking about her experience during a trip to Europe, and how happy she was to find a Turkish restaurant. While describing the reason for being happy for it, she uses the expression *hungry for twenty years*, which is clearly an exaggeration to express that they could not eat properly because they were in a foreign country and the food was quite different. Here, again, *like* seems to function as a focuser and mark the focus of the utterance.

4.6.2.6. Like to search for the appropriate expression

As a discourse marker, like can be also be used as the speaker is thinking about what to say next or is searching for the appropriate expression (Müller, 2005). According to Andersen (2001), it signals that 'I have something on my mind, but I don't know how to put it' (2001: 249). Such instances of *like* are usually accompanied by other hesitational markers such as filled and unfilled pauses. Examples of this function from the corpora examined in the present study are as in the following:

(38) <B> so things like that like factual issues cos I did something about vivisection <\B>
<A> mhm <\A>

<B> and then *like* .. er just creative writing just <X> fiction and stuff <\B>
NS Participant 10

In Example (38), the British interviewee uses *like* just before an unfilled pause and a filled pause, as she seems to search for the appropriate expression in the middle of an utterance.

(39) <A> I think being a farmer is like the ultimate job right </A>
<B> yeah in that movie yeah (erm) *like* . the .. what was its name . I can't remember the actors' name .. anyway the daughter finds out that there is a ghost in her room […] </B>
TR Participant 18

In like manner, in Example (39), the Turkish interviewee and the interviewer are talking about a movie that they have both seen before. As the interviewee wants to describe a scene in the movie, she wants to starts with the name of the actor, but cannot retrieve it. At the moment, *like* seems to mark a search for an expression along with a filled pause before and a very brief unfilled pause after it.

4.6.2.7. Like to restart/false start

The last function of *like*, i.e. *restart/false start*, that did not occur in the Turkish corpus and occurred only eight times in the LOCNEC is reported in quite a few studies (e.g. Schourup, 1985; Andersen, 2001; Müller, 2005). Instances that serve this function precede a restart that is "a point at which the present speaker stops an item under construction and recommences" (1985: 54). Examples of this function from the corpora examined in the present study are as in the following:

(40) <B> […] supposedly to give to another chocolate maker and said if you if you get <X> then I'll *like* I'll give you <XX> cash and then er Charlie says no no I'm not taking it it's not right gives it back to Willy Wanka Willy Wanka turns round says no no no you've passed the test I needed someone to have the chocolate factory when I get old it's yours <\B>
NS Participant 49

In Example (40), the British interviewee starts an utterance, then employs *like* and then restarts that same utterance. In this example, the utterance is cut off and resumes with the same words.

4.6.3. You know

The Turkish students who were interviewed for the Turkish corpus and their native English speaking peers from the LOCNEC employed *you know* for a variety of functions, with certain degrees of differences in the proportions of these functions in the two corpora. The overall findings are presented in Table 4.11 below.

Table 4.11. *Functions of you know in the Turkish Corpus and LOCNEC*

	Turkish Corpus		LOCNEC		Significance (X2/G2)
	n	%	n	%	
Textual					
Provide relevant background information	78	31.20	66	10.82	$p<.001$
Explanation/Explicitness	39	15.60	94	15.41	n.s.
Quotative	13	5.20	37	6.07	n.s.
Indicate topic change/digression	4	1.60	9	1.48	n.s.
Subtotal	134	53.60	206	33.77	$p<.001$
Interactional					
Confirmation seeker	34	13.60	124	20.33	$p<.05$
Mark lexical/content search	39	15.60	47	7.70	$p<.001$
Appeal for understanding	17	6.80	37	6.07	n.s.
Mark hesitation/uncertainty	13	5.20	121	19.84	$p<.001$
Reference to shared knowledge	9	3.60	9	1.48	n.s.
Restart/False start	0	0.00	9	1.48	n.s.
Self-repair	4	1.60	38	6.23	$p<.01$
Subtotal	116	46.40	385	63.11	$p<.001$
Unclassified Instances	0	0.00	19	3.11	
Total	250	100	610	100	

In terms of overall frequency, there was a considerably large difference in the frequencies of *you know* between the two corpora (250 times in the Turkish Corpus vs.

610 times in the LOCNEC). As for the functions these instances served in discourse, a total of 11 functions were identified in the corpora examined in this study. In the Turkish corpus, there were slightly more instances that functioned in the textual domain (about 54%) compared to those functioned in the interactional domain (about 46%). On the other hand, in the LOCNEC, there were more instances functioned in the interactional domain (about 64%) compared to those that functioned in the textual domain (about 34%). As is seen in Table 4.9, the Turkish students employed significantly more instances of *you know* that functioned in the textual domain compared to the British students. As for the interactional domain, the British students employed significantly more instances that functioned in this domain.

When the differences in individual functions are examined, the findings reveal significant differences, as is seen in Table 4.11, between the Turkish and British students when using *you know* to provide relevant background information ($p<.001$), seek confirmation ($p<.05$), mark lexical/content search ($p<.001$), mark hesitation/uncertainty ($p<.001$) and self-repair ($p<.01$). The Turkish students employed *you know* to provide relevant background information and mark lexical/content search significantly more frequently than the British students, whereas the British students used it to seek confirmation and mark hesitation/uncertainty significantly more frequently than the Turkish students. What all the functions reported in Table 4.11 mean and do in discourse are presented in the following sub-sections along with sample quotations from each corpus for every function revealed.

4.6.3.1. You know to provide relevant background information

In this function, *you know* can be used to introduce relevant background information, which can be in the form of a parenthetic comment (Östman, 1981; Erman, 1987). When serving this function, *you know* introduces "information that the speaker assumes the hearer needs to know so as to understand the prior message (Mei, 2012: 102). It is termed 'introducing relevant background information' since it is used to "mark inserts of parenthetic comments containing information that the speaker assumes the addressee needs to know in order to be able to follow" (Erman, 2001: 1344).

Examples of this function from the corpora examined in the present study are as in the following:

(41) <A> [I suppose you lost weight or <laughs> <\A>
<B> yeah we all lost weight *you know* I'm I was quite thin to start with so I didn't lose too much but . some people lost two or three stone <\B> NS Participant 50

In Example (41), the British interviewer comments on the interviewee's experience, and then, the interviewee says it is indeed the case stated in that comment, but provides some background information to indicate that it is not completely true for him.

(42) <B> […] people are speaking so differently from each other but what we have been learning is like the same but accents are quite different from each other . and the cultures quite different to . *you know* there are so many cultures in United States like Indian, Afro American and native Americans . Spanish yeah and each of them are so different and some of them hate each other some of them love each other </B>
<A> so that impressed you </A>
<B> yeah that impressed me a lot </B> TR Participant 31

Similarly, in Example (42), the Turkish interviewee is explaining why the United States impressed him as a country that he visited. He mainly points to the diversity of accents and cultures that exist in the US, and feels the need to provide some relevant background information regarding the extent of this diversity.

4.6.3.2. You know to provide explanation/explicitness

To provide explanation/explicitness is one of the frequent functions of *you know*, and introduces a clarification of some part of a previous statement (Erman, 1987). Clarification implies that the speaker wants to prevent a possible misunderstanding that the hear may have regarding an utterance (Müller, 2005). In other words, this function helps make a word or an expression more explicit for the hear's comprehension. Examples of this function from the corpora examined in the present study are as in the following:

(43) <B> [but they're not asking for money they're asking for chewing gum [and pens <\B>
<A> [really <\A>
<B> *you know* the pens that you push the top down <X> and they click <\B>
NS Participant 47

In Example (43), the British interviewee wants to clarify what she means by the word *pens* and employs *you know* to indicate that what comes next provides an explanation regarding what is preceded.

(44) <A> that's a little vague can you specify </A>
<B> yes especially when I go to the bridge *you know* Charles bridge (er) there was really a how can I say . how can I say gothic gothic atmosphere and I like that . there were statues a lot of statues on that bridge and it was like . I was feeling like they were comic to me . and would take me from there and . how can I say .. will take me from there and </B>
TR Participant 33

Likewise, in Example (44), the Turkish interviewee refers to a bridge while narrating her experience abroad, and wants to make explicit the name of the bridge that is marked by *you know*.

4.6.3.3. You know to introduce a quotation

In this function, *you know* marks "the transition from the speaker's own talk to a direct-speech report" (Redeker 1991: 1163). Within a turn, the speaker stops to provide a quotation from hypothetical conversation or a conversation that has actually happened before, and this is marked by *you know*. Examples of this function from the corpora examined in the present study are as in the following:

(45) <B> yeah well you you meet somebody who you think is completely new and you think oh *you know* I've met someone new and exciting or I don't know .. and it turns out they know all your friends from school and things <laughs> <\B> NS Participant 16

In Example (45), the British interviewee is talking about how it is like to live in a small city like Lancaster, and in a way, complains that everybody knows each other and

even though one thinks he/she meets someone new, it is not actually the case. She also wants to quote from her own hypothetical speech, and this quotation is marked by *you know*.

(46) <B> […] but the posture is the same so it's not about weight . and in the last one she is showing it off to her friends . *you know* I have a canvas picture on the wall *you know* that's me <laughing> so that's the story I think </B> TR Participant 36

In Example (46), there are two instances of *you know* that seem to have a quotative function employed by a Turkish interviewee. She is talking about the picture-story at the end of the interview, and narrates a possible quote from the woman in the pictures.

4.6.3.4. You know to indicate topic change/digression

In this function, *you know* indicates a topic change or a digression, i.e. a temporary departure from the main subject. Erman (2001) argues that it can form the boundary between topics. It is a digression if the speaker returns to the original topic before *you know*, but it becomes the new topic if he/she goes on the conversation with what is presented by *you know* (Müller, 2005). Examples of this function from the corpora examined in the present study are as in the following:

(47) <A> cos otherwise you die <laughs> <\A>
<B> yeah you die . so .. and I kind of realised <XX> couple of months that I could cope with anything because whatever I have to deal with I'm I'm gonna have to you know [you have to <\B>
<A> [mhm <\A>
<B> you survive and *you know* I think that's the most important lesson I've ever learnt that you erm .. you can cope with anything […] <\B> NS Participant 34

In Example (47), the British interviewee is talking about a past experience and how it was like, and then, she moves on to a further aspect, which is the lesson she learned from that experience. Again, it is not quite a complete shift from the topic of conversation, but a shift to a different issue within it.

(48) <A> <laughing> why didn't you choose something related to tourism then </A>
<B> my score was higher than that </B>
<A> oh okay you didn't wanna waste your score <laughing> </A>
<B> yes I didn't wanna waste it and also *you know* I want to graduate from this faculty so I can have a you know like (er) certain job . in the government so […] </B>
TR Participant 16

In a similar way, in Example (48), the Turkish interviewee firstly responds to a comment made by the interviewer, but then moves on to another aspect of the issue. This is not actually changing the topic completely, but rather, looking at the issue from a different perspective.

4.6.3.5. You know to seek confirmation

In this function, *you know* is used to seek the hearer's confirmation in conversation. In genres such as discussion in particular, individuals improve their utterances to make sure that there is a sufficient level of intelligibility when they observe any sign of misunderstanding or confusion from their addressee (Mauranen, 2006). In this sense, they often confirm whether a message has been communicated to the addressee as intended (Erman, 2001). In other words, *you know* can be used to "make sure that the listener has correctly understood specific references made in the text, usually to people but also to objects and other phenomena" (Erman, 2001: 1346). Examples of this function from the corpora examined in the present study are as in the following:

(49) <B> [I think the the way of life everything is really slow and laid back *you know* [and everything's <\B>
<A> [uhu they're relaxed <\A>
<B> yeah <\B> NS Participant 12

In Example (49), the British interviewee is talking about his journey to a South African country and describing the lifestyle there. As she describes it as *slow and laid back, s*he employs *you know* to a get a confirmation that the interviewer understands, which is indeed followed by a positive reaction by the interviewer.

(50) <B> […] that was the only place I saw that had that historical thing not in any other city I visited had that from Sakarya . I don't really . I'm not really familiar with things like that so it fascinated me . and it was interesting *you know* </B>
<A> uh-huh so what things did you learn from these monuments about the history of the area </A>
<B> apparently it was once a port for (er) […] </B> TR Participant 5

In Example (50), the Turkish interviewee is talking about a place that fascinated him, and at the end of the turn, he describes it as interesting and then employed *you know* presumably to confirm that the interviewer understands, after which the interviewer starts by expressing assent.

4.6.3.6. You know to mark lexical/content search

As a discourse marker, *you know* can also mark lexical/content search as the speaker is not certain about what to say. This usage is usually preceded or followed by a brief filled or unfilled pause.

(51) <B> [yeah survival <X> sort of like a a hundred crucial words or . [*you know* <\B>
<A> [phrases more phrases <\A>
<B> yeah yeah <\B> NS Participant 30

In Example (51), the British interviewee is talking about his desire to learn some basic French and refers to an amount of significant words and phrases, but the word *phrase* does not come to his mind apparently. He takes a short pause and employs *you know* while looking for the right word, and it is when the interviewer jumps in with that word.

(52) <B> […] I worked very good places like CNNTURK and Kanal D TGRT and I also had a (er) when I'm studying I go to the TRT 1 too for the *you know* . what was it <foreign> staj </ foreign> </B>
<A> internship </A>
<B> internship yes (er) so . at least I saw good you know places […] </B>
TR Participant 12

Likewise, in Example (52), the Turkish interviewee is talking about her experience as a journalist in different broadcasters. She also refers to her internship at the Turkish state television, but has difficulty in retrieving the word *internship* and uses *you know* in the meantime. After she says the Turkish equivalent of the word, the interviewer comes up with the English translation.

4.6.3.7. You know to appeal for understanding

In this function, the speaker uses *you know* to ask the hearer to understand the message anyway when he/she cannot find the right words. Müller (2005) argues that this function is related to 'lexical and content search' because in both cases there is an effort towards a search for what to say, but here, in *appeal for understanding*, the hearer is asked to get involved with the content. In other words, the speaker chooses to leave the message presented by *you know* to be fulfilled by the hearer (Mei, 2012). Examples of this function from the corpora examined in the present study are as in the following:

(53) <B> [I don't know if there is no .. erm .. I'm doing culture and communication which sort of involves .. films to a certain extent but er .. not a great deal <\B>
<A> [mhm <\A>
<B> [which is a shame cos I mean otherwise <begin_laughter> obviously <end_laughter> *you know* <\B>
<A> well you might . be able to find a course somewhere <\A>
<B> yeah <\B> NS Participant 20

In Example (53), the British interviewee leaves her utterance, in a way, incomplete when she uses *you know* at the end of the turn, waiting for the interviewer to imagine or understand what is meant by *otherwise*.

(54). <A> three days </A>
<B> yes . we stayed at a hostel . not too neat but *you know* </B>
<A> with too many people in the same room </A>
<B> yes . eight people </B> TR Participant 33

In Example (54), the Turkish interviewee is talking about her stay at a hostel, and how it was like. She says it was not too neat and uses *you know* possibly to indicate that the interviewer would understand what he means.

4.6.3.8. You know to mark hesitation/uncertainty

In this function, the speaker uses *you know* to buy more time to think about what should be said next because he/she seems hesitant or uncertain about it (Erman, 2001; Müller, 2005). As a hesitation marker, *you know* both helps the speaker gain more time, and also appeals to the hearer for patience or even for help (Mei, 2012). Examples of this function from the corpora examined in the present study are as in the following:

(55) <B> erm fifteen and sixteen <\B>
<A> oh yeah <\A>
<B> years old so so *you know* I I don't think they quite expected it to be like it was either <\B> NS Participant 42

In Example (55), the British interviewee is talking about a play to which he took his students while he was teaching in the past. Apparently, both he and his students did not expect the play as it was, and he is not quite sure how to express it as seems to be marked by *you know* and supported by the repetitive uses of *so* and *I* before and after *you know*. This uncertainty can be due to the interviewee's effort in looking for the right words, or trying to express himself appropriately.

(56) <A> so what things did you learn from these monuments about the history of the area </A>
<B> apparently it was once a port for (er) . international sea travelling trading .. I mean what did you learn I don't I didn't actually learn much well <laughing> I just *you know* visited . (er)</B>
<A> walking around the monuments </A>
<B> yeah </B>
<A> old places that fascinated you </A>
<B> yeah yeah </B> TR Participant 5

In Example (56), the Turkish interviewee is answering a question regarding his travel experience, and at the end of the turn, he seems hesitant about his choice of the

verb *visit* or whether it is sufficient to put it that way. This can be interpreted by both the use of *you know* before this choice, and a brief pause and the exclamation *er* expressing hesitation.

4.6.3.9. You know to make reference to shared knowledge

In this function, *you know* is used to focus the hearer on information which is actually shared. Here, shared means that it is something that has been previously talked about in conversation, or both the speaker and the hearer have a mutual experience of what is talked about. In such a context, *you know* is actually employed to make sure that the hearer has access such knowledge at that particular moment in discourse because it is important for the ongoing conversation (Müller, 2005). Examples of this function from the corpora examined in the present study are as in the following:

(57) <B> it's just for practicality's sake really <\B>
<A> mhm <\A>
<B> that it's like this but I mean I mean it's not attractive really you know all these big concrete buildings and like <\B>
<A> mhm <\A>
<B> they're building more near the library and things *you know* <\B>
<A> yes <\A>
<B> they're going up all the time and it's gonna end up a huge concrete jungle but <\B>
NS Participant 44

In Example (57), this time *you know* is employed at the end of the turn by a British interviewee, and since she studies in the same campus where the interviewer works, they must have a shared idea of the buildings located in the campus. In this particular instance, the interviewee refers to certain buildings near the library and uses the word *thing*s which should normally be vague, but considering that they are in the same campus, the interviewee makes reference to shared knowledge that is here marked by *you know*.

(58) <B> […] so I was on my own and it was a different experience you know the university was really different from what I . what I faced here <laughing> *you know* </laughing> our

university has a lot of exams presentations homework you know assignments but there I didn't do anything <laughing> […] </B> TR Participant 36

In Example (58), the Turkish interviewee is talking about her experience during an Erasmus experience at a European university, and compares her home university with that university. When referring to the study load at her own university, she employs *you know* at the beginning of the utterance, possibly to indicate that what comes next includes information both the interviewer and the interviewee know about.

4.6.3.10. You know to restart/false start

This is another function of *you know* that is reported in a number of studies (e.g. Erman, 2001; Müller, 2005), but occurred in the present study not very frequently. In the instances of *you know* that served this function, there is a prior message that is incomplete, and the following message usually continues where it is cut off, usually without any change (Mei, 2012). In such instances, the repetition usually spans only one or two words (Müller, 2005). Examples of this function from the corpora examined in the present study are as in the following:

(59) <A> and here it's just party and noise [and <\A>
<B> [publicity I hate it in France it was nice as well it's *you know* it's tradition and Christmas decorations come out in December in England […] <\B> NS Participant 32

In Example (59), the British interviewee starts an utterance, but then employs *you know* and starts the same utterance again. Thus, *you know* marks a restart here since there seems to be no correction or repair, but merely a repetition of one or two words, which in a way marks hesitation.

4.6.3.11. You know to self-repair

In this function, *you know* is used to rephrase the content, grammar or wording (Xue and Lei, 2016). For Erman (2001), it is a 'textual monitor' that signals repair, and occurs in places "where the speaker stops in mid-structure to make a restart" (2001:

1342). Examples of this function from the corpora examined in the present study are as in the following:

(60) <B> both they're both big cities <\B>
<A> .. is it . your favourite cities or <\A>
<B> yeah but I've I mean I've heard . a lot I've spoken to the fourth years who've just come back and . they were recommending .. whoever<?> <X> speak to it's always oh yeah go ahead go ahead cos it's always *you know* they loved it but I think I think I I'd like to go somewhere in the north <\B> NS Participant 14

In Example (60), the British student in this instance uses *you know* in the middle of an utterance, and continues the rest of the utterance with a different wording and organisation. It is slightly different from the previous example since this repair is more straightforward.

(61) <B> […] because lots of I watch lots of movies I always liked English and so I learn by myself to *you know* how to do this things and now I'm here I'm majoring (er) to English language (er) English language teaching […] </B> TR Participant 12

In like manner, in Example (61), the Turkish interviewee is talking about why she chose to study English language teaching, and she employs *you know* in the middle of an utterance as she figures out the rest of the utterance. Then she restarts the utterance, but with a difference organisation, which indicates a repair.

4.6.4. I mean

The Turkish students who were interviewed for the Turkish corpus and their native English speaking peers from the LOCNEC employed *I mean* for a variety of functions, with certain degrees of differences in the proportions of these functions in the two corpora. The overall findings are presented in Table 4.12 below.

In terms of overall frequency, there was a considerably large difference in the frequencies of *I mean* between the two corpora (87 times in the Turkish Corpus vs. 430 times in LOCNEC).

Table 4.12. *Functions of I mean in the Turkish Corpus and LOCNEC*

	Turkish Corpus		LOCNEC		Significance (X2/G2)
	n	%	n	%	
Textual					
Provide explanation/justification	17	19.54	57	13.26	n.s.
Reformulation (*in other words*)	4	4.60	29	6.74	n.s.
Explicitness/Clarification	7	8.05	48	11.16	n.s.
Exemplification	9	10.34	48	11.16	n.s.
Move back to the main point	4	4.60	10	2.33	n.s.
Subtotal	41	47.13	192	44.65	n.s.
Interpersonal					
Indicate Speaker Attitude	9	10.34	48	11.16	n.s.
Mark Certainty/Salience	0	0.00	9	2.09	n.s.
Subtotal	9	10.34	57	13.26	n.s.
Interactional					
Hesitation marker	16	18.39	0	0.00	$p<.001$
Self-Repair (Correcting grammar, word-choice, expressions)	4	4.60	38	8.84	n.s.
Restart (false start)	6	6.90	10	2.33	n.s.
Assumption-Correction	4	4.60	105	24.42	$p<.001$
Correcting Previous Statement	2	2.30	9	2.09	n.s.
Subtotal	32	36.78	162	37.67	n.s.
Unclassified Instances	5	5.75	19	4.42	
Total	87	100	430	100	

As for functions these instances served in discourse, a total of 12 functions were identified in the corpora examined in this study. Out of these functions, those in the textual domain constituted the majority, which were followed by functions in the interactional domain. The functions in the textual domain had the highest percentage in both the Turkish corpus and the LOCNEC.

When the differences in individual functions are examined, the findings reveal significant differences, as is seen in Table 4.12, between the Turkish and British students when using *I mean* to mark hesitation ($p<.001$), and correct hearer assumption

(p<.001). The Turkish students employed *I mean* to mark hesitation significantly more frequently than the British students, whereas the British students used it to correct hearer assumption significantly more frequently than the Turkish students. What all the functions reported in Table 4.12 mean and do in discourse are presented in the following sub-sections along with sample quotations from each corpus for every function revealed.

4.6.4.1. I mean to provide explanation/justification

In both corpora, *I mean* was employed to provide an explanation, background information or justification regarding an issue being talked about. Mei (2012) terms this function as justification "because *I mean* can be interpreted as 'I am saying this because' i.e. the upcoming utterance explains why the previous message has been said" (p. 77). She argues that the speaker's view and a further explanation of why he/she has stated it would more likely to win the hearer's agreement. Examples of this function from the corpora examined in the present study are as in the following:

(62) <B> and they were all .. <X> <\B>
<A> brilliant <laughs> <\A>
<B> brilliant it was just *I mean* it had you know you know Denzel Washington and Keanu Reeves and who would have thought you'd put Keanu Reeves <\B> NS Participant 22

In Example (62), the interviewer and interviewee were talking about a movie that they both thought was good. Then, the interviewee mentioned the actors who played in the movie so as to provide a justification why it was good, and this was marked by *I mean*.

(63) <A> and they are also Mediterranean people </A>
<B> yeah I know *I mean* how can I say it . I liked Portugal too I mean the people there are better even from Barcelona because they were so warm […] </B> TR Participant 29

In Example (63), the interviewee was talking about a travel experience in Spain and Portugal, and tried to justify why she liked the places she visited, or in other words provided an explanation.

4.6.4.2. I mean to reformulate

In this function, *I mean* can be regarded as '*in other words*', and Schiffrin (1987) considers it as a replacement repair. In this study, this function is coded as reformulation as it is a more common terminology. What speakers do here is basically repeat what is just said but in a slightly different wording, and this repetition should make it easier for the hearer to process the message. Examples of this function from the corpora examined in the present study are as in the following:

(64) <B> and of course we couldn't see anything cos the clouds were like lower than we were so *I mean* we were above the [clouds so we couldn't see anything <laughs> <\B> NS Participant 48

In Example (64), similarly, the British interviewee reformulated/rephrased what she just said (i.e. *being lower than the clouds* and *being above the clouds*). Here again, *I mean* seems to have been employed as *in other words*.

(65) <B> […] and when you see them in Vienna people are not harming them *I mean* they are protecting them . it was very beautiful the streets was so clean […] </B> TR Participant 8

In Example (65), the Turkish interviewee was talking about how the people of Vienna protected historical monuments/places, and used *I mean* to paraphrase what he just said in a different wording.

4.6.4.3. I mean to ensure explicitness/clarification

As a discourse marker, *I mean* can also be used to clarify or make content more explicit (Mei, 2012, Fernandez-Polo, 2014) as in *namely* or *that is*. This function constituted a considerable percentage among other functions (i.e. about 8% of all functions in the Turkish corpus, and about 11% of all functions in the LOCNEC).

Examples of this function from the corpora examined in the present study are as in the following:

(66) <A> you you don't want to do an M B A or something like [that or D B A or <laughs> <\A>
<B> [oh no . no no no <\B>
<A> [<XX> <\A>
<B> [I've had enough now *I mean* [I'm twenty-five so <\B> NS Participant 36

In Example (66), the British interviewer asked the interviewee whether she wanted to do, for instance, an MBA after completing her BA degree. She said she had had enough then, which can be regarded as an answer that one would expect from an older person, not from someone who is 22 or 23 years old, which is usually the age of graduation from BA degree. In other words, the interviewee's initial response may need some clarification, especially for someone who is not close. In this sense, the interviewee wanted to clarify it, and said she was already 25 years old, so apparently that is why she did not want to pursue further education. *I mean* marks this clarification just after the initial response.

(67) <B> okay I choose the third one . (er) it's not a city actually it's a district of Antalya . *I mean* Alanya (er) I have been to Alanya three or four times until nowadays […] </B>
TR Participant 10

In Example (67), at the beginning of the interview, the Turkish interviewee, after thinking about it for a few minutes, finally decided on the topic that he wanted to talk about. He chose the second topic, which reads as "*A country/city that you visited and that impressed you. Describe your visit and say why you found it particularly impressive*". However, what he said he wanted to talked about a district of a city, which apparently needs some clarification that is again marked by *I mean* right before it.

4.6.4.4. I mean to provide exemplification

This function of *I mean* relates to providing an example regarding an utterance that is usually rather broad. According to Mei (2012), it is called exemplification since

the following message functions as the backup of the preceding message, or as a further example for something rather broad. The instances of *I mean* in this function can be accompanied by expressions like *things like that* or *something like that*. Examples of this function from the corpora examined in the present study are as in the following:

(68) <A> have you ever been to Europe <\A>
<B> [erm <\B>
<A> [or<?> the continent well here it's Europe too I mean it it's funny because .. it's always as if England well Great Britain was not part of Europe when we're talking <\A>
<B> yeah .. <X> that's true .. erm . really I'm I'm not bothered I'd just like to travel <\B>
<A> mhm <\A>
<B> *I mean* I I'd love to go to New Zealand <\B>
<A> [New Zealand <\A>
<B> [which is somewhere I've never been <\B>
<A> mhm <\A>
<B> cos I've got a lot of relatives out there I've never seen <\B> NS Participant 17

In Example (68), *I mean* was again used, this time by a British interviewee, to indicate an example. He said that he liked traveling and provided an example, i.e. New Zealand.

(69) <B> […] I miss the life there in terms of . it was (er) when you . contrast here and there (er) if you don't think that euro is four liras <laughing> it was really cheaper there to live there *I mean* I was going to the market for buying something . I can't say to you how much I was buying it was just thirty euros I was just . I cannot buy even half of it in Turkey […] </B> TR Participant 29

Similarly, in Example (69), the Turkish interviewee was talking about an abroad travel experience and complaining about the low value of Turkish lira compared to euro, and then, wanted to provide an example, which was marked by *I mean*.

4.6.4.5. I mean to move back to the main point

As a discourse marker, *I mean* also indicates a return to the main point, or an issue that has been talked about just before. This function is sometimes referred to as resumption since the speaker picks up his/her topic that has been interrupted in the

previous discourse (Fernandez Polo, 2014; Mei, 2012). Examples of this function from the corpora examined in the present study are as in the following:

(70) <B> it's just brilliant and er and . I just like being off campus I think campus is a bit .. sort of too enclosed sometimes <\B>
<A> yes <\A>
<B> sometimes *I mean* having all the students is brilliant <\B>
<A> mhm <\A>
<B> but sometimes it's nice to get away from students <\B> NS Participant 27

In Example (70), the British interviewee was talking about his preference of living in campus over living in the city. He first said, also considering the previous context, that having other students in campus was brilliant, and then mentioned a downside. But at the end, he got back to the main point he wanted to communicate, which was marked by *I mean*.

(71) <A> so what things did you learn from these monuments about the history of the area </A>
<B> apparently it was once a port for (er) . international sea traveling trading .. *I mean* what did you learn I don't I didn't actually learn much well […] </B> TR Participant 5

In a similar vein, in Example (71), the Turkish interviewee was asked what he learned from his visits to historical monuments during a journey. He started his answer with some general information, but it was just not a direct answer to the question, or it was just not sufficient. Then, he wanted to go back to the main point, which was the question "what did you learn?", and employed *I mean* to indicate this shift.

4.6.4.6. I mean to indicate speaker attitude

In this function, *I mean* occurs with an evaluative adjective to indicate the speaker's attitude to something. It has been explored and reported in a quite a few studies and constituted most instances of *I mean* in different datasets (Brinton, 2007; Mei, 2012). Examples of this function from the corpora examined in the present study are as in the following:

(72) <B> in that sort of film *I mean* he was so .. he was such a baddie in the film but he came across so well <\B> NS Participant 22

Accordingly, in Example (72), the British interviewee expressed how she felt about the a character in the movie she was narrating, and *I mean* here signals that the upcoming message was the speaker's evaluation, judgement or attitude.

(73) <B> […] you don't see that in Sakarya . so I actually . my first impression was it's very nice *I mean* people are kind very . I mean okay it's not that bad . the city is not . so big […] </B> TR Participant 5

In a similar vein, the Turkish interviewee in Example (73) told the interviewee how he felt about the people living in the place he visited, and again *I mean* marks this personal evaluation.

4.6.4.7. I mean to mark certainty/salience

In this function, *I mean* can be used to indicate commitment to an idea, and includes an implicit claim that the speaker is certain about the information that is communicated (Fernandez Polo, 2014). In Schiffrin's terms (1987), this can be in the form of marking discourse salience in that the speaker creates stationary focus on a given topic, which indirectly claims that it is worth to pay attention to and encourage the hearer to focus on the items emphasised. Examples of this function from the corpora examined in the present study are as in the following:

(74) <B> and they show you this [I mean more or less <\B>
<A>[oh no must be awful <laughs> <\A>
<B> more or less but <\B>
<A> [yeah <\A>
<B> [*I mean* they do show him I mean there's a scene where they say they'll rack him so they get the horses and they're stretching him and it's pretty gruesome but <\B>
NS Participant 43

Consequently, in Example (74), the British interviewee was talking about a movie and particularly an actor in it. He told that this actor appeared at a particular scene, and

wanted to underline or emphasise that he was indeed shown in the movie. He employed *I mean* to stress what he just said.

4.6.4.8. I mean to mark hesitation

When *I mean* serves this function, it can occur with *you know* as well as filled and unfilled pauses. It is called a hesitation marker because, according to Mei (2012), here *I mean* accompanied with *you know* and/or pauses actually helps the speaker stall for time to figure out how to describe/put something in his/her mind. Examples of this function from the corpora examined in the present study are as in the following:

(75) <B> […] maybe she is saying to his friends as you can see I'm so beautiful my friend but in the (er) *I mean* real picture the man that was painting it she is not so much beautiful so maybe she wanted to change the I mean appearance . the appearance of herself in another's eyes I think . that I thought </B>
TR Participant 29

In Example (75), another Turkish interviewee employs *I mean* right after a filled pause, i.e. *er*, while trying to find the right expression, perhaps just to stall for time as she is processing the word-choice.

(76) <B> […] centre of the city there is a lot of this (er) Ottomans Romans all came here one time . <X> some . (er) very nice looking . buildings <X> (er) *I mean* . that was the only place I saw that had that historical thing […] </B> TR Participant 5

Again, in Example (76), the Turkish interviewee is talking about a city he visited, and seems to be hesitant about what to say, or how to verbalise what he thinks, which can be understood from a filled pause and an unfilled pause, and *I mean* is seen here in-between.

4.6.4.9. I mean to self-repair (correcting grammar, word-choice, expressions)

Self-repair is another function that occurred in the use of *I mean* in the both corpora, and it signals an upcoming modification or adjustment (Fox Tree and Schrock, 2002; Schiffrin, 1987). This function is similar to a restart/false start, but here there is an

actual modification of the prior utterance in the upcoming talk. Examples of this function from the corpora examined in the present study are as in the following:

(77) <A> biology that's quite different <\A>
<B> yeah up up to *I mean* in my first year I did English and biology and still couldn't decide […] <\B> NS Participant 24

In Example (77), the British interviewee is talking about her study of biology, and when the interviewer seems surprised, she starts providing an explanation. However, she feels that she had chosen the wrong preposition to refer to the time period in which she studied biology, and to correct that choice, she employs *I mean* and switches to a different prepositional phrase.

(78) <B> so when he speaks to er and he calls everyone brother [er like you can't do that to me brother <\B>
<A> [right yeah <\A>
<B> and he says thee and thou quite a lot as well but he says to er his parole officer I've done nothing wrong brother sir *I mean* and he he changes it <\B> NS Participant 49

In Example (78), the interviewee uses a term of address, i.e. *brother*, at the end of a quotation, but wants to correct it and change it to a different term, i.e. *sir*. Here, *I mean* is employed to mark this repair after the correction.

4.6.4.10. I mean to restart (false start)

Restart or false start was employed in the both corpora at a low level, and it indicates that the prior message is incomplete and followed by a complete one by another utterance marked by *I mean*. According to Mei (2012), this incomplete utterance can be due to the speaker's failure to organise his/her thoughts or the his/her decision to discontinue the current message. Examples of this function from the corpora examined in the present study are as in the following:

(79) <B> [we were there for about four year four days and erm . by the[i:] end *I mean* I only had a phrasebook by the[i:] end of the four days I was sort of you know I was the group's main spokesman […] <\B> NS Participant 26

In Example (79), the British interviweee talks about an experience and wants to refer to a time period at the beginning of the utterance, but then changes his mind in terms of organisation, and wants to restart and put that time period at the beginning of the next utterance. Here, *I mean* seems to mark that shift to the restart.

(80) <B> […] I think I wouldn't be that happy there because I'm really happy there you know like it is *I mean* it is a big city but not that big . transportation is more . you know easy […] </B> TR Particpant 16

In Example (80), the Turkish interviewee wants to provide an example, starts an utterance, but apparently has some difficulty in organising the rest of the utterance. Then, she uses *I mean* and restarts that utterance.

4.6.4.11. I mean to correct hearer assumption

This function that *I mean* was found to serve in the two corpora examined in this study is the one with the largest difference between the two. It was found in about 5% of all instances in the Turkish corpus, whereas in the LOCNEC this figure was around 25%, meaning that the British students considerably made use of this function of *I mean*.

This function is a subtype of repair, and is termed as assumption-correction since it is the hearer's potential assumption based on what is previously said that the speaker is trying to repair (Mei, 2012). In other words, *I mean* in this function helps the speaker to guide the hearer to interpret the previous message in the way the he/she intends. Examples of this function from the corpora examined in the present study are as in the following:

(81) <B> it was still quite eh . quite warm <\B>
<A> mhm <\A>

<B> but *I mean* it it did down at night <X> probably go down to about nought just above freezing <\B> NS Participant 50

In Example (81), the British interviewee is asked how the weather was like when he was in the Arctic. He first says it was not too bad and was quite warm, but then do not want the listener assume that the weather was always nice and well. Therefore, he uses *I mean* and then mentions the time it was not that warm; surely, the adversative conjunction *but* also contributes to the meaning in that what follows is the opposite of what precedes.

(82) <B> from re= first rehearsal to performance what happened was the[i:] examiner came and watched me .. eh doing a rehearsal and I picked a play that only had two characters in it so I wasn't working with loads and loads of people and we'd actually it was a bit of a cheat really because it was a bit we'd done before <\B>
<A> uhu <\A>
<B> and you know you had it it was kind of .. a rehearsed rehearsal but not word for word sort of thing so <\B>
<A> [yes <\A>
<B> [*I mean* obviously they know you're going to play it like that so [<XX> <\B>
NS Participant 9

In Example (82), the British interviewee is talking about a drama exam, or an audition, and says that they were supposed to play something they already knew before, and it was a bit of a cheating. But then, she does not want the interviewer to interpret this as they actually cheated the examiners. Rather, what she means, based on his last utterance, is that it was normal and the examiners also knew about it. This correction is also marked by *I mean* at the beginning of the turn.

(83) <B> […] I really like studying because of that I liked it but it was really really hard for us . *I mean* yes in here in Anadolu university the lessons are really so hard according to other . universities but there it was also hard […] </B> TR Participant 29

In Example (83), the Turkish interviewee is talking about her experience during Erasmus, and how difficult the courses she took there. But then, she wants correct a possible misunderstanding or assumption (i.e. the courses she takes in her own

university are easy), and uses *I mean* to provide that information in the following part of the utterance.

4.6.4.12. I mean to correcting previous statement

Another type of correction with which *I mean* seems to be employed is correcting a previous statement. This function resembles the previous function, i.e. assumption-correction, and self-repair, but here the speaker does neither really correct an assumption or a possible misunderstanding nor grammar or word-choice, but only a previous message he/she has has stated. Examples of this function from the corpora examined in the present study are as in the following:

(84) <A> would you recommend it to anyone or is it just . were you disappointed <\A>
<B> it's it's not not interesting <\B>
<A> it's not interesting <\A>
<B> *I mean* it's in a way it's much the same as as Britain <\B>
<A> yes <\A>
<B> well in in a way it's just a Western country you know <\B> NS Participant 25

In Example (84), the British interviewee is asked whether she would recommend a country she visited to anyone. She first says it is not interesting, which apparently surprises the interviewer, and then the interviewee employs *I mean* and wants to corrects or adjusts her initial answer.

(85) <B> it started with Enrique Iglesias in the because it was so popular during my high school years and everyone wants everyone was crazy about (er) Spanish actresses </B>
<A> he sings in English right </A>
<B> yes </B>
<A> mostly </A>
<B> he is not totally Spanish *I mean* he sings English songs too even sometimes Portuguese (er) […] </B> TR Participant 35

In Example (85), the Turkish interviewee is telling how she has become interested in the Spanish culture, and mentions the name of a Spanish singer. When the interviewer says whether that singer sings in English, she confirms and says that he is

actually half Spanish and sings English songs as well. Here, *I mean* is used just before this correction is provided.

4.6.5. Well

The Turkish students who were interviewed for the Turkish corpus and their native English speaking peers from the LOCNEC employed *well* for a variety of functions, with certain degrees of differences in the proportions of these functions in the two corpora. The overall findings are presented in Table 4.13 below.

Table 4.13. *Functions of well in the Turkish Corpus and LOCNEC*

	Turkish Corpus		LOCNEC		Significance (X2/G2)
	n	%	n	%	
Textual					
Mark continuation	6	11.11	20	3.76	$p<.05$
Opening of a discourse unit	5	9.25	10	1.88	$p<.01$
Indicate stages of a narrative	5	9.25	41	7.72	n.s.
Provide additional information	5	9.25	41	7.72	n.s.
Change of topic	1	1.85	20	3.76	n.s.
Mark quotation	0	0.00	31	5.83	n.s.
Subtotal	22	40.74	163	30.70	n.s.
Interpersonal					
Indicate an answer	19	35.19	153	28.81	n.s.
Evaluate a previous statement	4	7.40	20	3.76	n.s.
Express partial (dis)agreement, doubt, and contraposition	2	3.70	61	11.49	n.s.
Subtotal	25	46.30	234	44.07	n.s.
Interactional					
Search for the right phrase	4	7.40	10	1.88	n.s.
Turn-management	1	1.85	20	3.76	n.s.
Reformulation (rephrasing/correcting)	0	0.00	92	17.33	$p<.01$
Subtotal	5	9.25	122	22.98	$p<.05$
Unclassified Instances	2	3.70	10	1.88	
Total	54	100	531	100	

Since there was a considerably large difference in the frequencies of *well* between the both corpora (54 times in the Turkish Corpus vs. 531 times in LOCNEC), it may not seem very meaningful to make a significance comparison, but in overall, out of the three functional domains, the only significant difference was observed in the interactional domain. The Turkish students employed significantly less instances of *well* that functioned in the interactional domain compared to the British students. As for the textual domain, the Turkish students employed slightly more instances that functioned in this domain, while there was almost no difference regarding the interpersonal functions.

When the differences in individual functions are examined, the findings reveal significant differences, as is seen in Table 4.13, between the Turkish and British students when using *well* to mark continuation ($p<.05$), open a discourse unit ($p<.01$) and reformulate a previous utterance ($p<.01$). The Turkish students employed *well* to mark continuation and open a discourse unit significantly more frequently than the British students, whereas the British students used it to reformulate a previous statement significantly more frequently than the Turkish students. What all the functions reported in Table 4.13 mean and do in discourse are presented in the following sub-sections along with sample quotations from each corpus for every function revealed.

4.6.5.1. Well to mark continuation

To start with the textual domain, the function that appeared the most in the Turkish students' speech was to mark continuation. This function relates to the speaker's intention to say more about the topic he/she has just mentioned, and was reported by many scholars in the literature (e.g. Svartvik, 1980; Fischer, 1998; Schourup, 2001). It connects the preceding and following parts of texts together, and signals that something is going to follow (Schourup, 2001). Examples of this function from the corpora examined in the present study are as in the following:

(86) <A> [not just students <\A>

<B> [so it should be nice . that's right yeah . *well* where I'm living my aunt's just bought a house . er down behind the station and it won't be students cos it's a new housing area so I don't know who I'll be meeting but it'll be fun <\B> NS Participant 47

In Example (86), the British student compares how life is like in campus and in the city. As she complains that one usually meets students in the campus who mostly know each other, she argues that it is nice to meet not just students but other people as well. Then, after using *well* in the middle of the utterance, she moves on to talk about how it is like in his aunt's neighborhood where she lives at the time.

(87) <A> what are you planning do you plan to be a teacher </A>
<B> well . yes I do I'm going to courses right now . you know for KPSS exam . *well* I'm also studying for ALES . and you know YDS exam </B> TR Participant 31

In Example (87), the Turkish interviewee is asked about his future plans at some point in the interview. He wants to mention two things: the KPSS exam that is to work as a teacher, and the ALES exam that is to pursue a postgraduate education and/or work as a university lecturer/researcher. After telling the first one, he moves on to the second when he employs *well* and continues speaking.

4.6.5.2. Well to open a discourse unit

In this function, *well* can signal the development in discourse stage such as openings, and instances of *well* in such function signaling a new discourse stage or topic can occur initially or medially (Lam, 2009). Examples of this function from the corpora examined in the present study are as in the following:

(88) <A> right . so you're going to talk to me about a film aren't you <\A>
<B> right erm *well* the last film I watched was The Lion King <\B>
<A> [The Lion King <\A>
<B> [Disney's the Lion King <\B NS Participant 2

In Example (88), the British interviewee chooses to talk about a movie at the beginning of the interview. This new section, i.e. *moving from initial instructions and*

topic choice to talking about the movie, indicates the opening of a discourse unit, which is marked by *well*.

(89) <A> you are very fluent (er) we are just having an informal chat . we will begin with a topic I will give you actually three topics you will choose one start talking about it for maybe a few minutes and then we will have a conversation together </A>
<B> all right great </B>
<B> all right . *well* I have chosen second one a country that you visited and that impressed you </B> TR Participant 31

In Example (89), a similar situation can be observed in that the interviewee chooses a topic and starts talking about it after a few minutes of thinking at the beginning of the interview. Here again, *well* seems to indicate an opening.

4.6.5.3. Well to indicate stages of a narrative

This function was particularly observed when the interviewees were asked to describe a set of pictures in the form of a story. In other examples, it appeared in cases where the interviewees were trying to narrate a previous experience regarding the topic of conversation. In the following two examples, *well* was employed in this respect.

(90) <A> it's a story and I'd like you to tell me what's going on in the pictures <\A>
<B> ... <laughs> right the[i:] artist erm would you like me to sort of speech bubble it <XX> sort of<?> put [words into their mouths <laughs> <\B>
<A> [oh just .. the way you want to <\A>
<B> erm *well* in the first picture the[i:] artist is is erm drawing the[i:] the sitter and then er <\B> NS Participant 37

(91) <A> that's great <laughing> because students usually go for the third one </A>
<B> (er) really . okay . it's easy <laughing> (er) .. when I was in Prague I decided to go Vienna . but (er) me and my two friends (er) . bought our tickets to go there . and (er) we barely arranged this trip because we are also work we were working and we didn't (er) we didn't have any (er) free days to go there . *well* we made it and we bought our tickets . and (er) last day my other friends decided not to go there and […] </B> TR Participant 27

In Example (90), *well* marks the beginning of the narrative as the British interviewee is asked to narrate or describe the four pictures within the picture-description task at the end of the interview. On the other hand, in Example (91), *well* occurs as the Turkish interviewee is narrating a travel experience, and *well* marks the next scene in her narrative.

4.6.5.4. Well to provide additional information

In this function, *well* introduces explanations and additional information regarding the preceding discourse. In other words, it can be described as an introducer of an explanatory comment (Lam, 2009). Examples of this function from the corpora examined in the present study are as in the following:

(92) <B> [<X> ... the[i:] Opera House is really nice <\B>
<A> did you go in there .. in [the <\A>
<B> [yeah yeah *well* I di= I didn't actually see an opera but I sort of [just went into it <\B> NS Participant 25

In Example (92), the interviewee was talking about the Opera House and how nice it was, and then, the interviewer asked whether she had been in there before or not. It was a simple yes-no question which the interviewee answered as "yeah yeah", but then he felt the need to provide additional information probably because she thought the interviewer thought she saw an opera when she went into the Opera House. At this point, *well* seems to mark that an additional information was to be provided.

(93) <A> and where are you from </A>
<B> from Istanbul </B>
<A> from Istanbul okay </A>
<B> *well* that's not my father's hometown but I was born in Istanbul and raised in Istanbul so I'm from Istanbul </B> TR Participant 8

In Example (93), the interviewer asked the interviewee's hometown, and he said he was from Istanbul, but then provided an extra information regarding that his

ancestors were not really from Istanbul, again probably because he wanted to only mean that he grew up and lived in Istanbul to avoid any misunderstandings.

4.6.5.5. Well to change the topic

In this function, *well* marks a topic shift; this can be a change to an entirely different topic or a new aspect of the same topic. According to Lam (2009), this role of *well* is similar to punctuation marks in writing that divides words into clauses and sentences.

(94) <B> cos that was part of the problem <X> I was on my own I just felt .. erm <\B>
<A> how about the[i:] other girls I mean <\A>
<B> *well* the stupid thing was that when this girl got violent <\B>
<A> [she got violent .. oh no <\A> NS Participant 46

In Example (94), the interviewee was talking about his apartment and the people with whom she shared it. Suddenly, she started to share a previous experience with one of her mates. This shift was again marked by *well*.

(95) <B> okay I think (erm) second question is good for me because I remember the most beautiful city for me I think it is Istanbul I went there for once and I stayed there just one day *well* I think it is impressive because of the weather the sea the people . there were lots of people crowded city and it's a really big city […] </B> TR Participant 15

In Example (95), after having thought about which topic to choose at the beginning of the interview, the interviewee finally decided on the topic and briefly described it, and then moved on to talking about specifics regarding why the city she chose impressed her. Here, *well* seems to mark the boundary in a shift from a general, straightforward entrance to the topic to more specific aspects.

4.6.5.6. Well to mark quotation

Similar to *like*, *well* can also function as marking a quotation. In Svartvik's terminology (1980: 175), it functions "as a signal indicating the beginning of direct speech, parallel to that of quotation marks in writing" (1980:175). The quotative use of

well has been mentioned by a number of scholars (e.g. Svartvik, 1980; James, 1983, cited in Müller, 2005, p. 114; Schiffrin, 1987; Jucker, 1993). Moreover, Schiffrin argues that "*well* marks the orientation shifts created by reported speech, i.e. talk whose original time, place, and possibly author is concurrent with the ongoing conversation" (1987: 124). Examples of this function from the corpora examined in the present study are as in the following:

(96) <B> I'm sorry I'm sorry <X> you can't <X> you're from South Africa yeah I know I know but but <X> yeah <X> er <XXX> but people say so what are your favourite films and I say A Clockwork Orange Reservoir Dogs Natural Born Killers that kind of thing and they say and what else and I say *well* Willy Wanka and the chocolate factory and the Wizard of Oz <\B> NS Participant 49

In Example (96), the interviewee was talking about an imaginative conversation, and employed *well* to quote from his own speech in that conversation. Most of the examples in the LOCNEC in this function was in the form of quoting one's own speech in such context.

(97) <B> with not preparing anything just walk in and say right yeah<?> [*well* tell me about <\B>
<A> [well today <laughs> <\A>
<B> a film you've seen in English [kind of thing or get them to work in groups you know just <XX> text <\B> NS Participant 39

Similarly, in Example (97), the interviewee was narrative a previous experience about her teaching experience, and she quoted from her speech as she hypothetically entered the classroom and instructed the students what to do.

4.6.5.7. Well to indicate an answer (direct/indirect)

Indicating or prefacing an answer is probably one of the most dominant functions of *well*, which appeared quite frequent in both the Turkish Corpus and LOCNEC. The responsive use of *well* that function in the interpersonal domain is usually the most common among other functions, which is in line with observations reported in previous

studies (Lakoff 1973; Schiffrin 1987). In this sense, as a responsive signal, it is usually associated with dis-preferred responses such as disagreements and criticisms (Lam 2006), though it is also found to preface direct answers and follow-up responses (e.g. Schourup 2001; Aijmer and Simon-Vandenbergen 2003; Müller 2005; Lam 2008). Examples of this function from the corpora examined in the present study are as in the following:

(98) <A> do you often go rock climbing in the Lake district when you're here [because it's it's so near <\A>
<B> [mm .. *well* yeah it's getting there is a bit of a problem I'm I'm. a member of the[i:] mountaineering club <\B> NS Participant 50

Example (98) also seems similar to Example (100), in that it prefaced an answer that was not quite direct accompanied by a further explanation.

(99) <A> do you ever plan to go back there </A>
<B> *well* I can but if I had lots of money to spend . you know I can but . you know if I don't have money . why why should I go there </B> TR Participant 31

In Example (99), the interviewee was asked a question and started his answer with *well.* This is probably because there was no simple and easy answer for the question, considering that he provided various conditions just after the short positive answer to the yes/no question.

(100) <A> okay . so did you . write any specific programs </A>
<B> (er) *well* here is the thing I actually program the things I want and I would use so not for a commercial use but I want . worked in <X> <X> I don't know how to say it <X> . people pronunce it differently . I think you may be familiar with it </B> TR Participant 5

Similarly, in Example (100), another Turkish interviewee started his answer to a yes/no question with *well*. This answer was particularly indirect although the question was direct. This could be due to the fact that the interviewee's answer would not meet the interviewer's expectation while asking the question, and he needed to provide

additional/further information so as to indicate why that expectation of a simple answer could not be met.

4.6.5.8. Well to evaluate a previous statement

This function is also referred to as a mitigator of some sort of confrontation (Svartvik, 1980), In this regard, Smith and Jucker (2000: 216) argues that "a typical function of *well* is to downgrade a claim". Accordingly, this function of *well* serves to make the upcoming uterance less face-threatenning (Müller, 2005). In the present study, in some of the instances, *well* also marked an evaluation of a previous comment or statement made by the other interactant. As we see in the two examples below, the interviewees employed *well* at the beginning of their evaluation of the interviewer's statement.

(101) <A> oh that's nice and how long are you going to stay there <\A>
<B> in Geneva erm I guess twelve to fifteen months something like that depending on . [<X> it takes me to do <X> <\B>
<A> [so we won't w= we're not we're not going to see you for <\A>
<B> *well* <X> it's possible <X> I'm gonna be back in England in September .. but I'll be down in Oxford but I could probably make it back up here <\B> NS Participant 30

In Example (101), the interviewer made a comment about the upcoming long-term travel of the interviewee and said they would not see him for a while. Then, the interviewee provided a further comment by saying that it was possible.

(102) <A> with maps you know you could zoom in zoom out . see the world so he just (er) tried to find his own village using google earth spent like hundreds of hours . and at the end he actually did meet his mom . and it's inspired from a real story . at the end of the movie they actually . show the moment that the step mom and the real mom meet together and they cry and hug . that's very impressive <laughing> </A>
<B> <laughing> *well* that's good </B> TR Participant 13

In Example (102), the interviewer described the storyline of a movie the he saw recently, and then the interviewee provided an evaluation that was marked with *well*.

4.6.5.9. Well to express partial (dis)agreement, doubt, and contraposition

As a discourse marker, *well* is also frequently associated with partial (dis)agreements regarding a previous statement. More specifically, *well* seems to indicate partial agreement, doubt, (partial) disagreement, and contraposition. According to Aijmer (2011), *well* can function as a downtoner and softener before an opinion expressing an opinion or feeling such as disagreement or disapproval. Examples of this function from the corpora examined in the present study are as in the following:

(103) <A> yeah so erm so your mother is Scot= <\A>
<B> *well* she's half Scottish yeah <\B> NS Participant 43

In Example (103), the interviewer made inference based on the interviwee's previous statements, but the interviewee stated his partial agreement with this inference which was again marked by *well* at the beginning.

(104) <A> perhaps it makes it more exciting maybe right <laughing> </A>
<B> *well* maybe but you know it was kind of ridiculous because (er) the man in the movie was . had a sight problem . and you know the other woman who also had the same problem was attracted to him because they had the same health issue and I think that represents how we choose our . maids <laughing> it's kind of ridiculous </B>
TR Participant 1

In Example (104), on the other hand, the interviewer stated an opinion about a characteristic of a movie that he found exciting, but then the interviewee thought it was ridiculous rather than being exciting. Her disagreement was marked by *well* at the beginning of her turn to indicate that she was not completely in agreement with the previous statement, and a disapproval was to be followed.

4.6.5.10. Well to search for the right phrase

In this function, *well* indicates that the speaker is in search of the right phrase (Müller, 2005), and indicates the hearer that he/she needs some time for processing, or

even needs some help. Examples of this function from the corpora examined in the present study are as in the following:

(105) <B> and I thought well especially since I saw it *well* . over a year ago and I can still remember a lot of it so clearly <\B> NS Participant 33

In Example (105), the interviewee was talking about a previous experience and needed to specify a time period, but said *well* and paused shortly before briefly thinking how much time she should approximately state.

(106) <B> yeah you got a lot of fun while learning it and you don't get bored at all . in Turkey this situation is very different from other countries . and what can I say more about Slovakia . *well* . </B>
<A> so what did you do during your visits you just visited the city or </A>
TR Participant 8

In Example (106), the interviewee compared the life in Turkey, Slovakia and other countries, and at the end of the turn he felt like he should say more about Slovakia, and after a short pause, he briefly thought about what else to say, which was marked by *well*, and afterwards, the interviewer jumped in and asked another question.

4.6.5.11. Well to manage turn transition

In this function, *well* can signal speakers' desire to take control of the conversational floor, in that there are two types: floor-holding and turn-taking. In floor-holding, the speaker is interrupted while holding the conversational floor, and then may need to reinstate their speaker role as someone else wishes to take over. On the other hand, in turn-taking, the speaker can use *well* to seize the right to voice their opinion from the current speaker when he/she is in the hearer role. Examples of this function from the corpora examined in the present study are as in the following:

(107) <A> well yeah well I mean <laughs> if you d= if you have to do that for six weeks sometimes you m= might get hungry in the morning [and <X> <\A>

<B> [*well* I usually<?> had a chocolate bar <laughs> and you have breakfast about nine o'clock <\B> NS Participant 32

In Example (107), as the interviewer was talking and apparently had more to say since he used an additive conjunction at the end of his turn, the interviewee wanted to state her own opinion of the issue and took the conversational turn by using *well*.

(108) <B> what's going on . all right </B>
[…]
<B> *well* should I name the characters </B>
<A> no I mean just . </A> TR Participant 8

In Example (108), the interviewee agreed that he understood the instruction for the picture-description task through the end of the interview. After a planning time for about five seconds, the interviewee had a question about the task, and during that silence, he took the floor by using *well* and asked his question.

4.6.5.12. Well to reformulate (rephrase/correct)

In this function, *well* can be an editing marker for self-correction (Svartvik 1980) and be used to rephrase something that you have just said. Examples of this function from the corpora examined in the present study are as in the following:

(109) <A> for foreign people <\A>
<B> for foreign . I think or or may= maybe for everybody *well* maybe for certain Kenyans as well cos you've got certain very rich areas and very poor areas and . […] <\B> NS Participant 46

In Example (109), the interviewee needed to rephrase an utterance she just produced; after using the word “everybody”, she wanted to limit the scope to “certain people”. For this purpose, she used *well* to indicate that she needed to rephrase it.

(110) <A> <XX> quite scary <\A>
<B> yes I know .. but normally there's there's a big queue so *well* or a couple of people or whatever so <\B> NS Participant 32

Similarly, in Example (110), another interviewee employed *well* while he rephrased an utterance.

4.7. Relationship Between Speaking Fluency and the Use of Discourse Markers

In order to analyse the relationship between the Turkish students' speaking fluency and their use of discourse markers, the analysis of their speech in terms of temporal and perceived indicators was conducted, and is presented in this section.

4.7.1. Utterance fluency

In order to determine the Turkish students' utterance fluency, various temporal indicators such as speech rate (no. of syllables and words per minute), articulation rate, mean length of runs, mean pause duration, and mean duration of syllables were calculated. The descriptive statistics regarding these indicators are presented in Table 4.14.

Table 4.14. *Descriptive statistics regarding the temporal indicators of speaking fluency in Turkish students' speech*

	SPM	WPM	AR	MLR	MPD	MDS
N	50.00	50.00	50.00	50.00	50.00	50.00
Mean	158.35	118.26	3.55	8.45	0.71	0.29
Std. Deviation	24.58	18.06	0.56	3.38	0.12	0.05
Minimum	112.39	84.19	2.51	3.66	0.50	0.20
Maximum	212.72	157.29	4.95	20.07	0.97	0.40

SPM: Syllables per minute, WPM: Words per minute, AR: Articulation rate
MLR: Mean length of runs, MPD: Mean pause duration, MDS: Mean duration of syllables

As is seen in Table 4.14, the Turkish students produced 158.35 syllables per minute in average, with 112.39 syllables per minute (TR Participant 40) as the minimum and 212.72 syllables per minute (TR Participant 36) as the maximum value. In another indicator of speech rate, they produced 118.26 words per minute in average, with 84.19 words per minute (TR Participant 40) as the minimum and 157.29 words words per minute (TR Participant 36) as the maximum value. As for articulation rate,

i.e. total number of syllables divided by phonation time, the Turkish students produced 3.55 syllables per second, with 2.51 syllables per second (TR Participant 2) as the minimum and 4.95 syllables per second (TR Participant 41) as the maximum value. In another measure, which is the mean length of runs, i.e. total number of syllables divided by the total number of pauses, they produced 8.45 syllables per run, with 3.66 syllables per run (TR Participant 40) as the minimum and 20.07 syllables per run (TR Participant 6) as the maximum values. In terms of mean pause duration, i.e. the total duration of pauses divided by the total number of pauses, the Turkish students paused 0.71 seconds in average, with 0.50 seconds as the minimum (TR Participant 12) and 0.97 seconds (TR Participant 40) as the maximum value. Lastly, in mean duration of syllables, i.e. phonation time divided by the total number of syllables, the Turkish students spent 0.29 seconds per syllable, with 0.20 seconds as the minimum (TR Participant 41) and 0.40 seconds (TR Participant 2) as the maximum values.

For speech rate, the average speech rate of native speakers tend to range from 120 to 260 words per minute (e.g. Richards 1983; Suenobu et al. 1986). It seems to be a wide range since it takes different communication situations into account. Tauroza and Allison (1990) provides a chart for the range of speech rate in British English speech in interviews. In this chart, the following rates are given: fast (above 250 wpm), moderately fast (210 to 250 wpm), average (160 to 210 wpm), moderately slow (120 to 160 wpm), and slow (below 120 wpm). In this respect, the Turkish students in the present study can be regarded as slow with a mean speech rate of 118.26, and even the participant with the highest wpm (157.29) can be categorised as having nearly average speech rate. As for the mean-length of runs (MLR), Grosjean and Deschamps (1975, cited in Götz, 2013) reported that the MLR of a native speaker ranges between 7.42 syllables (in descriptions) and 14.85 syllables (in interviews). In non-native speech, Hincks (2008) revealed that Swedish learners produced runs with a mean of 12.59 syllables per phrase in English, and this is 24% shorter than when they speak in their L1. Compared to both native speakers and Swedish learners, the Turkish students in the present study can be regarded as less fluent in terms of MLR with a mean of 8.45 syllables per minute.

4.7.2. Perceived fluency

In order to determine the Turkish students' perceived fluency, a native English speaker rated the Turkish participants who were interviewed for the Turkish corpus based on a 9-point scale, i.e. 1 being extremely dysfluent to 9 being extremely fluent. The descriptive statistics regarding this rating are presented in Table 4.15.

Table 4.15. *Descriptive statistics regarding the perceived rating of speaking fluency in Turkish students' speech*

	NS Fluency Rating
N	50.00
Mean	6.62
Std. Deviation	1.10
Minimum	5.00
Maximum	9.00

As is seen in Table 4.15, the mean rating score of the participants' perceived fluency was found to be 6.48 out of 9. The participants who were rated as the least fluent (TR Participants 10, 11, 17, 22, 24, 27, 30) received 5 out of 9, while those who were rated as the most fluent (TR Participants 7, 41, 42, 48) received 9 out of 9. If 5 out of 9 is taken as moderately fluent, the Turkish students' perceived fluency by the native speaker can be regarded as being slightly above moderately fluent.

4.7.3. Correlation between utterance/perceived fluency and frequencies of discourse markers

In order to reveal the relationship between the Turkish students' use of discourse markers, and their utterance and perceived fluency, the bivariate Pearson correlation was conducted by using SPSS. A correlation matrix yielded from this analysis is presented in Table 4.16.

Starting with speech rate, the correlation analysis revealed significant moderate correlations between the total normalised frequency of the five discourse markers employed by the Turkish students, and the syllables ($r=.35$, $p<.05$) and words ($r=.40$, $p<.05$) they produced per minute.

Table 4.16. *Correlation matrix of the relationship between the Turkish students' discourse marker use, and their utterance and perceived fluency*

		DMs	SPM	WPM	AR	MLR	MPD	MDS	NS
DMs	Pearson Correlation	1	.358*	.409*	.339*	.116	-.123	-.338*	.466**
	Sig. (2-tailed)		.028	.011	.037	.489	.462	.038	.003
	N	50	50	50	50	50	50	50	50
SPM	Pearson Correlation		1	.979**	.617**	.578**	-.626**	-.606**	.407*
	Sig. (2-tailed)			.000	.000	.000	.000	.000	.011
	N		50	50	50	50	50	50	50
WPM	Pearson Correlation			1	.597**	.580**	-.627**	-.587**	.445**
	Sig. (2-tailed)				.000	.000	.000	.000	.005
	N			50	50	50	50	50	50
AR	Pearson Correlation				1	-.139	.088	-.984**	.521**
	Sig. (2-tailed)					.404	.601	.000	.001
	N				50	50	50	50	50
MLR	Pearson Correlation					1	-.663**	.147	.148
	Sig. (2-tailed)						.000	.379	.376
	N					50	50	50	50
MPD	Pearson Correlation						1	-.104	-.050
	Sig. (2-tailed)							.535	.767
	N						50	50	50
MDS	Pearson Correlation							1	-.475**
	Sig. (2-tailed)								.003
	N							50	50
NS	Pearson Correlation								1
	Sig. (2-tailed)								
	N								50

*. Correlation is significant at the 0.05 level (2-tailed).
**. Correlation is significant at the 0.01 level (2-tailed).
DMs: Total normalised frequency of the five discourse markers, SPM: Syllables per minute,
WPM: Words per minute, AR: Articulation rate, MLR: Mean length of runs, MPD: Mean pause duration,
MDS: Mean duration of syllables, NS: Native speaker rating

For articulation rate, there was a significant moderate correlation between the students' discourse marker use and the number of syllables they produced per minute ($r=.34$, $p<.05$). A weak correlation was revealed between the students' discourse marker use, and their mean length of runs ($r=.11$), while a weak negative correlation was found in their mean pause durations ($r=-.35$). As for mean duration of syllables, a significant and negative moderate correlation was found between the students' discourse marker use, and the mean duration they spent per syllable ($r=.-34$, $p<.05$). Lastly, a significant moderate correlation was revealed between the students' discourse marker use, and their perceived fluency as rated by a native English speaker ($r=.47$, $p<.01$).

5. DISCUSSION

In this section, the findings revealed in the study are discussed with regard to the research questions addressed above. References are made to the relevant literature on various aspects of the use of discourse markers in different contexts.

5.1. Summary and Discussion of the Findings Related to the First Research Question

With regard to the first research question, which reads as *Are there any differences between Turkish and British university students in terms of the frequency and function of discourse markers they employ?*, the Turkish and British university students' use of discourse markers was compared in terms of frequency and functions.

5.1.1. Frequency of the discourse markers in the Turkish corpus and LOCNEC

The findings show that both groups employed *so* with the highest frequency among the five discourse markers examined, and it was followed by *like* and *you know*, respectively. The Turkish students used *well* with the least frequency, while the discourse marker with the lowest frequency in the British students' speech was *I mean*. In comparison to the British students, the Turkish students significantly underused four out of the five discourse markers examined, namely *so*, *you know*, *I mean* and *well*. However, in terms of normalised frequency, they used *like* as a discourse marker almost at the same level with their native English speaking peers (i.e. 4.99 times vs. 5.34 times per 1,000 words, respectively), although this slight different was found to be significant when the frequencies of the individual students were compared between the two corpora.

These findings do not mostly overlap with other studies in the literature, particularly regarding the use of *well* by non-native speakers. Buysse compared the use of *well*, among other discourse markers, in non-native and native speech in two studies (2010, 2015). In Buysse (2010), he compared the use of Flemish and British university students within a comparable interview corpus. The Flemish university students used *well* extremely more frequently than their native speaker peers did, and naturally than

Turkish students did in this study (7.40 times vs. 0.90 times per 1,000 words, respectively). Moreover, in Buysse (2015), he examined *well* in the speech of Dutch, French, German, Spanish and Chinese learners of English as well as in that of British students. He reported that all learner groups, except Chinese learners, significantly overused *well* compared to the British students, and the Chinese learners who significantly underused this discourse marker had a normalised frequency that is quite close to that of the Turkish students in this study (0.76 times vs. 0.90 times per 1,000 words, respectively). Chinese learners' underuse of *well* was also observed in several other studies (e.g. Fung and Carter, 2007; Liao, 2009). Likewise, the two other studies that also focus on Turkish speakers of English reveal similar findings regarding the use of *well*. Aşık and Cephe (2013) reported that the Turkish university students studying ELT used *well* 0.20 times per 1,000 words in in-class presentations, while Zorluel Özer and Okan (2018) found that the Turkish EFL teachers used *well* 0.11 times per 1,000 words in their lectures, and this was significantly less frequent than the use of native English teachers. These normalised frequencies of *well* in other studies on Turkish speakers of English are even lower than the one reported in this study (i.e. 0.90 times per 1,000 words). Accordingly, Turkish individuals seem to underuse *well* while speaking English in various contexts including informal interviews, in-class presentations and lectures.

General arguments can be developed for possible reasons behind the underuse of discourse markers by non-native speakers, but for a specific discourse marker (e.g. *well* in this case), certain perspectives can be taken while explaining a variation in use. The L1 equivalent of that discourse marker can be considered in this respect. Buysse (2010) explains the Flemish students' overuse of *well* by arguing that its Dutch equivalents (*wel, nou (ja), nu (ja)*) was extremely frequent in spoken Dutch, and for this reason, the Flemish students had a positive, or at least neutral, attitude towards this discourse marker, which was reflected in their spoken English. As for Turkish, *well* does not seem to have an agreed equivalent; while Yilmaz (1994) referred to *şey* as the Turkish equivalent of *well*, Tekin (2015) argued that *işte* had more functions in common with *well* and used *işte* as its Turkish equivalent. In these two studies, the Turkish equivalents of *well* were argued to have functions that are largely similar to those in English and

they are also commonly used in Turkish. This means that Buysse's (2010) interpretation may not be fully valid for the Turkish students since the Turkish students significantly underused *well* although its Turkish equivalents were common, and shared largely similar functions. In another study, which was a case study based on interviews with a Turkish speaker of English, Polat (2011) was not able to detect even one occurrence of *well* as a discourse marker in the interviews she conducted with the Turkish participant over a one-year period. It can thus be argued that Turkish students' underuse of *well* might be attributed to their lack of the necessary knowledge about how *well* may function in spoken English. This lack of knowledge may be due to the multifunctional nature of *well* as it has 12 functions (as is revealed in this study) that operate at textual, interpersonal and interactional levels, and interpersonal and interactional functions constitute half of these 12 functions. The argument that discourse marker functions that operate at non-textual planes/levels such as interactional and interpersonal are expected later than textual functions (Hays, 1992) may be the case in the use of *well* by Turkish speakers of English. Additionally, although the Turkish equivalents of *well* are said to have functions that are similar to *well*, these equivalents (i.e. *işte* and *şey*) may not be as multifunctional as *well* when evaluated individually.

In the current study, *I mean* and *you know* were other discourse markers with significant differences between the two corpora. While the Turkish students in this study employed *I mean* 1.60 times and *you know* 4.18 times per 1,000 words, the Turkish students in Aşık and Cephe (2013) used them 0.29 times and 0.37 times per 1,000 words, and the Turkish teachers in Zorluel Özer and Okan (2018) used them 0.32 times and 0.45 times per 1,000 words, respectively. In Aşık and Cephe (2013), *I mean* and *you know* occurred slightly more frequently in the native speaker reference corpus, whereas in Zorluel Özer and Okan (2018) *I mean* occurred in native teachers' lectures almost with the same frequency as the Turkish teachers while there was a significant difference in the use of *you know*. The differences between these three studies may be due to the nature of the data used in these studies; while Aşık and Cephe (2013) used oral presentations of students on a topic specific to a subject area course, Zorluel Özer and Okan (2018) recorded lectures of teachers, but this study conducted informal interviews

with students. So, their data seems less dialogic, and more planned and monologic compared to the data in this study, which could, in a way, explain the differences.

You know and/or *I mean* were also reported to be underused by Flemish, Dutch, French, German, Spanish, Chinese, Cantonese and Japanese learners (Fung and Carter, 2007; Liao, 2009; Buysse, 2010, 2017; Shimada, 2014), and also sometimes significantly underused when compared to native speakers. Two arguments can be produced to interpret this underuse by individuals from various L1 backgrounds including Turkish. The first one is that these two discourse markers might be problematic in spoken language for learners of English no matter what their L1 background is. Non-native speakers as a group might be learning these two discourse markers quite late since the aforementioned studies including the present study reported a significant underuse. Secondly, informal interviews, as a genre, which is the data used in this study, might not require the use of these two markers, but this does not seem to be the case, because the British students in the LOCNEC made use of them quite frequently. Another option is the lack of exposure to the discourse marker functions of these lexical items since such functions are more likely to occur in informal speech contexts, which is less likely the case in an EFL classroom.

With regard to *like*, the British students were found to use it slightly more frequently than the Turkish students, but this difference was significant according to the Mann-Whitney U test although it was not significant based on the log-likelihood statistic. This inconsistency is presumably due to the large extent of variation in the Turkish data; 11 Turkish students never used *like* as a discourse marker and nine students used it only once, which accumulates to 40% of the Turkish students. Therefore, it can be argued that the Turkish students slightly underused *like* compared to their British peers. In studies that focused on learners with various L1 backgrounds, learners underused *like* compared to native speakers, mostly significantly (Müller, 2005; Buysse, 2010; Shimada, 2014; Gilquin, 2016). However, the difference between Turkish and British students in this study seems to be the closest one reported in relevant studies that conduct a similar comparison. The use of *like* is far more limited by the Flemish students in Buysse (2010), the Japanese students in Shimada (2014) and 11 L1 background groups as a whole (Bulgarian, Chinese, Dutch, French, German, Greek,

Italian, Japanese, Polish, Spanish and Swedish that constitute the LINDSEI corpus) in Gilquin, 2016). The finding that the Turkish students approximate to native speaker performance to a larger extent than other L1 background groups can be due to the Turkish equivalent, i.e. *falan*, being argued to have a strong resemblance to *like* in English. In a cross-linguistic comparison of the functions of *falan* in Turkish and *like* in English, Tekin (2015) concludes that *falan* has strong resemblance to the discourse marker *like* in English, which can be argued to play a role in Turkish EFL learners' use of *like* while speaking English.

Lastly, *so* was the most frequent of the discourse markers examined in both the Turkish corpus and LOCNEC. However, it was still used significantly more frequently by the British students in the LOCNEC. The results reported in the literature are somewhat inconsistent. While Flemish and Japanese learners were reported to significantly overuse it when compared to their native English speaking peers (Buysse, 2010, 2012; Shimada, 2014), Cantonese and German learners were found to use it less frequently than the native speakers (Fung and Carter, 2007; Müller, 2005). To some extent this inconsistency might be due to the nature of the data examined in these studies; the former two employs informal interviews as the primary dataset (i.e. the same in the present study), while the latter two employs role-playing tasks and retelling tasks based on a silent movie, respectively. If we consider the studies that most corresponds to the data used in the present study, the Turkish learners (18.46 times per 1,000 words) seem not to have deviated too much in the frequency of *so* compared to Flemish (12.10 times per 1,000 words) and Japanese (20.68 times per 1,000 words) students while interacting in informal interviews. On the other hand, the Turkish students employed almost all functions of *so* that the native speakers used, and significantly underused only one out of 10 functions. Therefore, it can be argued that the Turkish students had a considerable degree of familiarity with *so* as a discourse marker in terms of both its frequency and functions. This can be because *so* has a number of textual functions (e.g. indicate a result and introduce a summary) that may be more explicit to learners, and it seems more likely that EFL learners may be exposed to such functions quite frequently in their classroom environment, at least more likely than its other functions, and even the functions of other discourse markers.

5.1.2. Functions of the discourse markers in the Turkish corpus and LOCNEC

As for the functions of the discourse markers examined, certain differences were also observed between the Turkish corpus and the LOCNEC. To begin with *so*, which is the most frequent discourse marker in overall, a total of 10 functions were identified, and the British students employed it in interpersonal functions more often than the Turkish students did, while the Turkish students used it in textual functions more often than their British peers did. In the use of *like*, a total of seven functions were identified, and most of its instances functioned in the textual domain. When the differences between the two corpora are concerned, the the Turkish students employed *like* to introduce an explanation and provide a quotation significantly more frequently than the British students, whereas the British students used it to introduce an example significantly more frequently than the Turkish students. As for *you know*, there were a total of 11 functions identified in the two corpora. The Turkish corpus included a slightly higher proportion of textual functions unlike the LOCNEC that included more instances that functioned in the interactional domain. When it comes to *I mean*, there were a total of 12 functions that operated mostly in the textual domain, followed by the interactional domain. The Turkish students employed *I mean* to mark hesitation significantly more frequently than the British students, whereas the British students used it to correct hearer assumption significantly more frequently than the Turkish students. Lastly, in the use of *well*, a total of 12 functions were identified in the two corpora examined. Regarding the functional domains, the Turkish students had a significantly lower proportion of interactional functions compared to their British peers.

In overall, the Turkish students employed the functions that served mostly in the textual domain. In the literature, interpersonal and interactional functions are usually reported to be used more frequently by native speakers. For instance, Mei (2012) reported that *I mean* fulfilling an assumption-correction function was considerably more often in the speech of British speakers than Chinese EFL learners, whereas resultative and conclusive *so* were used more frequently by Dutch and German EFL learners (Müller, 2005; Buysse, 2012). This may be because interpersonal/interactional functions are associated with informality, and thus, students may have felt reluctant to use them in

an interview setting simply because they lack pragmatic and register awareness (Buysse, 2010). If informality is in play, then the level of acquaintance between the interviewee and the interviewer can also be influential; when the two persons are familiar with each other, it is more likely that the interaction in-between can be more informal. On the other hand, some of the functions, e.g. to mark hesitation, have a somewhat different circumstance. In the case of this function, it was employed by British students by means of *you know* more often than the Turkish students, whereas it is the vice-versa in the Turkish students' use of *I mean* to mark hesitation. In other words, the two groups used two different discourse markers to fulfil the same function. However, in Mei (2012), this function almost never occurred with *I mean* in the speech of British students, but it did occur quite frequently in the speech of the Chinese students. Similarly, in study on the functions of *I mean* and *you know* in native speaker discourse, Fox Tree and Schrock (2002) state that hesitations often co-occur with *you know*, and suggests that *you know* is employed "when speakers are having extra trouble expressing themselves, to encourage addressees to infer the intentions" (p. 738), but do not attribute such a function related to hesitation to *I mean*. Moreover, *you know* was also reported to have an editing function that is sometimes highlighted by hesitation in Buysse (2017), and this function was found to be employed by native speakers much more often compared to four different L1 groups. Therefore, it can be argued that some discourse functions are achieved by non-native speakers with discourse markers different from what native speakers would use to fulfil those functions. Another issue that can be in play is whether the functions of L1 equivalents of a discourse marker are similar to its functions in the L2, for which cross-linguistic studies might provide valuables insights.

A further issue that might also have been in play is the familiarity between the interviewer and the interviewees/participants during the data gathering process. In the compilation of the Turkish corpus, the interviewer/research was more familiar/ acquainted with some of the participants, while not that much with others. Since certain functions of the discourse markers are more related to informal contexts, the participants might have wanted to be more formal during the interviews, which may have influenced their choice and use of certain discourse markers and their functions.

Though, it can also be valid for the British student corpus, or LOCNEC, in which not in all interviews the interviewer and the interviewee were acquainted with each other.

5.2. Summary and Discussion of the Findings Related to the Second Research Question

With regard to the first research question, which reads as *Are there any differences between first year and fourth year Turkish university students in English language teaching (ELT) in terms of the frequency of discourse markers they employ?*, the first and fourth (i.e. final) year university students' use of discourse markers was compared in terms of frequency. In both groups, *so* was the most frequent discourse marker, and *well* was the least frequent. However, the first-year Turkish students employed the five discourse markers examined in the present study more frequently than the fourth-year students. Moreover, this difference was significant in four of these five markers, i.e. *so*, *like*, *you know* and *I mean*, in favour of the first-year students. The only exception that did not reveal a significant difference in-between was *well* that was employed again more frequently by the first-year students, but the difference was not significant.

This research question was asked with an aim to compare beginning and graduating EFL teacher candidates' use of discourse markers so as to have a preliminary idea about their performance as future teachers. This can be regarded as important because one of the arguments that are stated with respect to non-native speakers' lack of awareness regarding the use of discourse markers is that learners are not provided with such lexical items in EFL classrooms, or teachers do not raise awareness regarding the varying functions of such items. In this respect, EFL teacher candidates' performance can give us clues about how they would deal with this issue because if they themselves are not aware of different functions of discourse markers, it is not likely that they will bring such items and their functions to the attention of their students in the future. Therefore, it is surprising that the fourth-year Turkish students who were about to graduate at the time of the interviews, and some of whom probably started teaching at schools across Turkey as this thesis is being written, showed a poorer performance in the use of discourse markers.

One aspect that might be related to the fourth-year or graduating students' poorer performance is that during their four-year education they also gain some professional experiences through both their academic courses and school internships, and near the end of their education they get ready for the work life. Such experience and development may have encouraged these students to adopt a more formal attitude during the interviews although these interviews were informal. Additionally, since also discussed in various parts of this work, certain discourse markers and their functions are more related to informal contexts, these students who possibly had a more formal attitude due to their increased professional experience may have avoided such lexical items accordingly. Although this is only an educated guess or speculation, it may also be an issue that interacted with fourth-year students' performance.

Even though the fourth-year students' performance might be influenced by their increased professional experience, various cross-sectional studies that examine different language areas/skills in ELT university students also show that fourth-year students do not perform any better than those who just completed their first year. Durmuşoğlu Köse and Yüksel (2013) found that second-year ELT students' vocabulary size was larger than fourth year students, while Şener (2015) reported a steady decrease throughout the four-year period, and that first-year students performed far better than their fourth year peers in terms of vocabulary size. As for writing, Tömen (2016) reported that there was almost no difference between the essay scores of first and fourth year ELT students. The common finding that graduating students perform poorer or at least not better than beginning students in different studies including the present study is surprising because the language of instruction is naturally English in four-year ELT programs in which language-skills courses and teaching methodology courses are taught along with elective linguistics and literature courses. In other words, students are exposed to a quite large amount of target language in different modes. However, most of the skills courses are embedded in the first and second years although they also read and write in English in most of the academic courses in the rest of their education.

In brief, the present study revealed that first-year ELT university students used the five discourse markers examined in the study significantly more frequently than their fourth-year peers. Yet, this finding does not seem to be unexpected, considering that the

literature reports poorer or at least not better performance of fourth-year students compared to first-years in vocabulary and writing. In a way, this could be due to the nature of the ELT programs in Turkey, but fourth-year students' increased experience near the end of their education may also be an influential issue.

5.3. Summary and Discussion of the Findings Related to the Third Research Question

With regard to the third research question, which reads as *Is there a relationship between Turkish university students' use of discourse markers, and their utterance fluency and perceived fluency?*, the relationship between the Turkish students' use of discourse markers, and their utterance and perceived fluency was examined. According to the results, the Turkish students' use of discourse markers had positive, moderate and significant relationships with their speech rate (i.e. both in syllables and words produced per minute), articulation rate, and perceived fluency ratings by a native speaker, whereas it had a positive, weak and non-significant correlation with their mean length of runs. As for negative relationships, the Turkish students' use of discourse markers had a negative moderate and significant correlation with their mean duration of syllables, and a negative and non-significant correlation with their mean pause durations. This means that the Turkish students produced more syllables and words per minute, produced more syllables within their phonation time and are perceived as more fluent, as they employed discourse markers more frequently in their speech. Likewise, as they employed such lexical items more frequently, they spent less time to produce a syllable in average, and paused for shorter durations while speaking. In other words, although the relationship is mostly at a moderate level, as they used discourse markers more frequently, they had better indicators of temporal fluency, and are perceived as more fluent.

Olynyk et al. (1987) who argued that the use of discourse markers, which they called speech markers, can be seen as a positive contributor to listeners' perception of non-native speakers' fluency. However, Götz (2013) reported that the use of a high proportion of discourse markers did not have a positive impact on the temporal fluency performance of native and non-native speakers. She also looked at perceived fluency

based on native speaker ratings, but she did not specifically include the use of discourse markers in that analysis. The difference can be explained by the methodology that Götze adopted; she calculated a 'fluenceme' score that is a combination of various temporal fluency indicators. Thus, her calculation seems more comprehensive in terms of temporal fluency, whereas the temporal fluency indicators are simpler and more straightforward in this study. Moreover, she did not specifically focus on the use of discourse markers and perceived fluency. However, considering that the correlation reported in this study is at moderate level and Götze (2013) did not report a positive impact, the relationship between the use of discourse markers and speech fluency might not be as strong as it is argued in the literature (e.g. Aijmer, 2002; Hellermann and Vergun, 2007). Nevertheless, fluency of speakers can vary depending on communicative situations, and this study focused only on a specific situation, i.e. informal interviews. Therefore, the relationship between the use of discourse markers and different dimensions of speech fluency should be examined in a variety of communicative situations and genres to better understand the degree and nature of such relationship.

6. CONCLUSION

6.1. Concluding Remarks

This study aimed to investigate the use of discourse markers by Turkish and native English-speaking British university students, and to determine whether discourse marker use is in fact related to speaking fluency. The students' use of discourse markers was examined in informal interviews, which is a relatively more dialogic genre compared to those adopted in previous research on Turkish EFL learners.

The results indicated that the Turkish students significantly underused four out of the five discourse markers examined, namely *so*, *you know*, *I mean* and *well*, in comparison to the British students. However, in terms of normalised frequency, they used *like* as a discourse marker almost at the same level with their native English speaking peers, although this slight difference was found to be significant when the frequencies of the individual students were compared between the two corpora. As for the functions of these discourse markers, the Turkish students employed the functions that served mostly in the textual domain, while the British students made use of a higher proportion of interpersonal and interactional functions compared to the Turkish students.

In addition, among the Turkish students who were interviewed to compile the Turkish corpus, those who were in their first year were compared to their fourth-year peers in terms of the frequency of the discourse markers they employed in the interviews. The first-year ELT students used the five discourse markers examined in the study significantly more frequently than their fourth-year peers.

Moreover, the Turkish students' use of discourse markers had positive, moderate and significant relationships with their speech rate (i.e. both in syllables and words produced per minute), articulation rate, and perceived fluency ratings by a native speaker, whereas it had a positive, weak and non- significant correlation with their mean length of runs. Besides, the Turkish students' use of discourse markers had a negative, moderate and significant correlation with their mean duration of syllables, and a negative and non-significant correlation with their mean pause durations. In other words, the Turkish students produced more syllables and words per minute, produced

more syllables within their phonation time and are perceived as more fluent, as they employed discourse markers more frequently in their speech. In a similar vein, as they employed discourse markers more frequently, they spent less time to produce a syllable in average, and paused for shorter durations while speaking. However, the relationship was at moderate level, and thus, the connection between the use of discourse markers and speech fluency might not be as strong as it is argued in the literature (e.g. Aijmer, 2002; Hellermann and Vergun, 2007). After all, there are other, and perhaps more significant, variables that contribute to individuals' utterance and perceived fluency. Moreover, the ideal extent of discourse marker use might not be as obvious as it is thought because as discussed above informality seems to be an important aspect of discourse marker use in speech, and there are a wide range of situations and contexts that can be placed on different spots across the informality continuum. To sum up, although the relationship is mostly at a moderate level, as the Turkish used discourse markers more frequently, they had better indicators of temporal fluency, and are perceived as more fluent.

Consequently, it can be concluded that the Turkish ELT university students differed in the use of discourse markers from their native English speaking British peers in terms of both frequency and functions, in that they underused the discourse markers examined in the present study, and used these markers in textual functions in a higher proportion and interpersonal/interactional functions in a lower proportion compared to the British students. Yet, there was a moderate relationship between the Turkish students' use of these discourse markers, and their utterance and perceived fluency.

6.2. Implications

6.2.1. Pedagogical implications

Pragmatic competence is a crucial component of communicative competence, and a key element that is argued to play a significant role in the pragmatic competence and spoken interaction of a speaker is discourse markers (Müller, 2005; Mullan, 2010). Considering that foreign language learners should have pragmatic competence in the target language so as to be competent users of that language, learners are expected to be

proficient in the use of certain elements and their functions in spoken discourse. As one of these elements, discourse markers are lexical items that individuals use to create textual coherence and express their feelings and views (Carter and McCarthy, 2006), and are argued to be pragmatically indispensable in spoken discourse (Mei, 2012). Since discourse markers help achieve a variety of functions in conversation, and are considerably frequent in native speaker discourse, it can be argued that they do have a pedagogical value. The fact that many studies including the present one report that non-native speakers of English differ in their use of discourse markers than native speakers in terms of frequency and function demonstrates this value even more. As the focus of this study is on Turkish students, then discourse markers can also be of value for English language instruction in Turkey as well.

With regard to the pedagogical neglect of discourse features, Fung and Carter (2007) emphasise that "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" (p. 433). In other words, discourse use seems to be neglected in foreign language classrooms. They also argue that the adverb, adjective, and noun meanings of *well* are covered in most language classrooms, but its pragmatic usages in spoken English to fulfil various discourse functions such as changing to topic and managing turns are rarely focused on (ibid). Thus, there seems to be a general neglect of the knowledge of discourse markers in foreign language teaching in spite of its importance in native speaker discourse and non-native speakers' poor performance in this respect (Romero Trillo 2002; Müller 2005).

Aijmer (2002) argues that non-native speakers' incorrect use or underuse of discourse markers may result in misunderstandings. Moreover, they can be perceived as dysfluent when they perform non-target-like use of discourse markers (Hellermann and Vergun, 2007). Consequently, discourse markers that are used by native speakers effortlessly is necessary for learners to express themselves in a fluent and confident way in that language (Sankoff et al., 1997). If they are indeed neglected in foreign language instruction, which seems to be the case based on non-native speaker performance, then

ways to include them in instruction should be sought after. For instance, Fung and Carter (2007) suggest that the language awareness-based III (Illustration-Interaction-Induction) approach, which is proposed by McCarthy and Carter (1995), implemented by means of activities such as language observation, problem-solving, and cross-language comparisons can be instructive in making explicit the meaning and use of discourse markers. Similarly, language samples from daily conversations in native speaker discourse can also be employed to demonstrate the use of discourse markers in natural conversation (Hellermann and Vergun, 2007). On the other hand, language teaching materials and coursebooks should also touch upon such features of spoken English with examples. As Lam (2009) reports, significant discrepancies exist between teaching materials and examples from naturally-occurring speech. Therefore, course curricula and materials should be evaluated in this respect.

Above all, it is EFL teachers who should raise learners' awareness regarding the role and function of discourse markers in native speaker discourse. In a way, they set a model for learners as they are the only source of L2 input at school in a foreign language learning environment. Considering that the present study was conducted with Turkish preservice EFL teachers who demonstrated a somewhat poor performance regarding the use of discourse markers, it can be argued that EFL teacher training programs should equip student teachers with the knowledge related to role and function of various features of spoken English like discourse markers. Based on the finding that the first-year ELT students performed better the their fourth-year peers, it can be suggested that the skills-based courses in the first year of ELT programs can be spread through the four-year period to a certain extent.

6.2.2. Methodological implications

This study is believed to contribute to the literature as revealing the discourse marker use of Turkish students in comparison to native English students, and demonstrating the relationship between discourse marker use and different indicators of speech fluency. It is methodologically different particularly from previous research on Turkish students' use of discourse markers in terms of two aspects: examining a more dialogic genre and limiting the analysis to the most frequent and most widely studied

discourse markers. Firstly, previous studies (Aşık and Cephe, 2013; Zorluel Özer and Okan, 2018) focused on planned and more dialogic speech like in-class student presentations or teachers' lectures. However, certain functions, especially those that operate in interpersonal and interactional domains, are less likely to occur in such genres since they involve less interaction. Indeed, the data gathered in the present study yielded a wide variety of discourse markers in both the Turkish corpus and LOCNEC. Secondly, both the aforementioned studies had a very broad definition of discourse markers and focused on almost all lexical items that could be found to have a discourse marker function in their data. However, limiting the scope to certain discourse markers, most preferably to the most frequent and most commonly studied ones, can yield more meaningful results. As a matter of fact, the present study went beyond merely reporting the frequency information and broad functional categories for discourse markers, and analysed individual functions of each marker examined. In this way, not only the frequency-related differences between Turkish and native English students but also the variations in the functions that the discourse markers they used fulfilled in their speech could be revealed.

6.3. Suggestions for Further Research

Based on the results and implications of the study, several suggestions can be offered for further research. These suggestions can be gathered around genre, research design, in-service teachers, possible influential factors, L1 effect, functional analysis, positional analysis, and teaching discourse markers:

- The genre focused in the present study is informal interviews. Although it is relatively more dialogic and interactive compared to recorded in-class student presentations and lectures, it may not reveal certain functions of discourse markers since one of the interactants is in the interviewer role and the other in the interviewee role. Therefore, corpora consisting of more informal communication situations can be compiled to examine the use of discourse markers in such contexts. For instance, Turkish students' personal encounters or chats in English during their Erasmus study-abroad experience can be

recorded for a corpus project, and such data would possibly provide more natural and interactive conversations to be examined,

- A cross-sectional research design was adopted in the present study, in that first and fourth-year ELT students were interviewed and compared. A similar comparison can also be conducted across different proficiency levels for EFL learners in general. Yet, a longitudinal study would demonstrate how learners' acquisition of discourse markers in an EFL context develops over a long period of time,

- The present study focused on pre-service EFL teachers, or ELT university students. However, a similar attempt on in-service teachers could also provide insights in terms of revealing their use and competence regarding spoken discourse markers. It could be of significance, considering that it is teachers who are expected to raise students' awareness in this respect,

- Possible influential factors can also be examined in a further study. Such factors may include age, gender, personality, study-abroad experience and context. There are certain attempts to examine the role of age and gender in the use of discourse markers in native speakers, in that they were indeed found to vary based these variables (e.g. Laserna, M. C., Seih, Y. and Pennebaker, J. W., 2014). However, such variables can also be examined for non-native speakers and EFL contexts,

- As is referred to while discussing some of the findings, the L1 equivalents of L2 discourse markers may influence L2 learners' use of discourse markers. For instance, there might not be an exact equivalent, or it may have different functions than in L2. Therefore, comparative or cross-linguistics studies can be conducted, comparing the use and functions of discourse markers in the first and target languages,

- Certain functional studies are available in the literature, especially the functions of *you know* and *I mean* in comparison. Such efforts can be extended

to other discourse markers, and the nuance differences in the same functions of different discourse markers should be examined so as to define what factors affect native speakers choice of one discourse marker over the other to fulfil the same function in discourse,

- In this study, the positions of the discourse markers within utterances were not examined since it was not an aim or research question in the beginning. Nevertheless, the position of a discourse marker may be related to the functions it fulfils in discourse. Therefore, their positions can also be included in such a study,

- As is reported in the literature and the present study, non-native speakers, although they are mostly presumably upper-intermediate to advanced learners, demonstrate a poor performance regarding spoken discourse markers. Consequently, one of the pedagogical implications is that such elements of spoken discourse should be included in EFL instruction, curricula, materials and coursebooks. Yet, how they should be taught or presented to students in an EFL classroom to enable their acquisition of these elements can be focused in experimental studies. Such efforts would provide important insights to teachers, teacher trainers, and material and curriculum developers.

REFERENCES

Abraham, W. (1991). Discourse particles in German: how does their illocutive force come about?. In A. W. Abraham (Ed.), *Discourse particles: Descriptive and theoretical investigations on the logical, syntactic and pragmatic properties of discourse particles in German*, (pp. 203-252). Amsterdam: John Benjamins.

Aijmer, K. (2002). *English discourse particles: Evidence from a corpus*. Amsterdam: John Benjamins Publishing.

Aijmer, K. (2004). Pragmatic markers in spoken interlanguage. *Nordic Journal of English Studies*, 3(1), 173-190.

Aijmer, K. (2005). Evaluation and pragmatic markers. In E. Tognini-Bonelli & G. Del Lungo Camiciotti (Eds.) *Strategies in academic discourse,* (pp. 83-96). Amsterdam: John Benjamins.

Aijmer, K. (2011). Well I'm not sure I think… The use of *well* by non-native speakers. *International Journal of Corpus Linguistics*, 16(2), 231-254.

Aijmer, K. and Simon-Vandenbergen, A.M. (2011). Pragmatic markers. In J. Zienkowski, J. Östman, J. Verschueren (Eds.), *Discursive Pragmatics* (pp. 223–247). Philadelphia: John Benjamins.

Andersen, E. S., Brizuela, M., DuPuy, B. and Gonnerman, L. (1999). Cross-linguistic evidence for the early acquisition of discourse markers as register variables. *Journal of Pragmatics*, 31(10), 1339-1351.

Andersen, G. (1997). They like wanna see like how we talk and all that: The use of *like* and other pragmatic markers in London teenage speech. In U.-B. Kotsinas, A. B. Stenström, and A. M. Karlsson (Eds.), *Ungdomsspråk i Norden* (pp. 82–95). Stockholm: MINS.

Andersen, G. (1998). The pragmatic marker *like* from a relevance-theoretic perspective. In A. H. Jucker and Y. Ziv (Eds.), *Discourse Markers: Description and Theory* (pp. 147–170). Amsterdam: John Benjamins.

Andersen, G. (2001). *Pragmatic markers and sociolinguistic variation: A relevance-theoretic approach to the language of adolescents*. Amsterdam: John Benjamins.

Andersson, L. G. and Trudgill, P. (1990). Bad language. Penguin.

Aşık, A. and Cephe, P. T. (2013). Discourse markers and spoken English: Nonnative use in the Turkish EFL setting. *English Language Teaching*, 6(12), 144-155.

Bachman, L. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.

Bachman, L. F. and Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests* (Vol. 1). Oxford: Oxford University Press.

Bazzanella, C. (1990). Phatic connectives as interactional cues in contemporary spoken Italian. *Journal of Pragmatics,* 14, 629-647.

Blakemore, D. 1987. *Semantic constraints on relevance*. Oxford: Blackwell.

Blakemore, D. (1988). 'So' as a constraint on relevance. In R. M. Kempson (Ed.), *Mental representations: The interface between language and reality*, (pp. 183–195). New York: Cambridge University Press.

Blakemore, D. 1992. *Understanding utterances: An introduction to pragmatics*. Oxford: Blackwell.

Blakemore, D. (2002). *Relevance and linguistic meaning: The semantics and pragmatics of discourse markers.* Cambridge: New York: Cambridge University Press.

Blyth, C., Recktenwald, S. and Wang, J. (1990). I'm like,"say what?!": A new quotative in American oral narrative. *American Speech*, 65(3), 215-227.

Boersma, P. and Weenink, D. (2007). Praat: doing phonetics by computer (Version 6.0.37)[Computer program]. http://www.fon.hum.uva.nl/praat/ (Access Date: 04.02.2018).

Boxer, D. (1993). Complaints as positive strategies: what the learner needs to know. *TESOL Quarterly*, 27, 277–299.

Brinton, L. J. 1990. The development of discourse markers in English. In J. Fisiak (Ed.), Historical linguistics and philology, (pp. 45-71). Berlin: Mouton de Gruyter.

Brinton, L. J. (1996). *Pragmatic markers in English: Grammaticalization and discourse functions.* Berlin; New York: Mouton de Gruyter.

Brinton, L. J. (1998). 'The flowers are lovely; only, they have no scent.': The evolution of a pragmatic marker in English. In R. Borgmeier, H. Grabes & A. H. Jucker

(Eds.), *Anglistentag 1997 Giessen. Proceedings*, (pp. 9-33). Tier: Wissenschaftlicher Verlag.

Brinton, L. J. (2007). The development of I mean: Implications for the study of historical pragmatics. In S. M. Fitzmaurice and I. Taavitsainen, *Methods in Historical Pragmatics* (pp. 37-74). Berlin, New York: Walter de Gruyter.

Bu, J. (2013). A study of the acquisition of discourse markers by Chinese learners of English. *International Journal of English Studies*, 13(1), 29-50.

Buysse, L. (2010). Discourse markers in the English of Flemish university students. In I. Witzcak-Plisiecka (Ed.), *Pragmatic Perspectives on Language and Linguistics, Vol. 1: Speech Actions in Theory and Applied Studies*, (pp. 461-484). Cambridge Scholars Publishing, Newcastle upon Tyne.

Buysse, L. (2012). *So* as a multifunctional discourse marker in native and learner speech. *Journal of Pragmatics*, *44*(13), 1764-1782.

Buysse, L. (2015). 'Well it's not very ideal…' The pragmatic marker *well* in learner English. *Intercultural Pragmatics,* 12(1), 59-89.

Buysse, L. (2017). The pragmatic marker *you know* in learner Englishes. *Journal of Pragmatics*, 121, 40-57.

Carter, R. and McCarthy, M. (2006). *Cambridge grammar of English: A comprehensive guide to spoken and written English grammar and usage*. Cambridge: Cambridge University Press.

Crystal, D. (1988). "Another look at, well, you know..." *English Today, 13*, 47–49.

Dailey−O'Cain, J. (2000). The sociolinguistic distribution of and attitudes toward focuser like and quotative like. *Journal of Sociolinguistics*, 4(1), 60-80.

De Cock, S. (2004). Preferred sequences of words in NS and NNS speech. *Belgian Journal of English Language and Literature* (BELL), 2, 225–246.

De Jong, N. and Wempe, T. (2008). Praat Script Syllable Nuclei v2. https://sites.google.com/site/speechrate/Home/praat-script-syllable-nuclei-v2 (Access Date: 08.02.2018).

De Jong, N., Steinel, M. P., Florijn, A., Schoonen, R. and Hulstijn, J. H. (2013). Linguistic skills and speaking fluency in a second language. *Applied Psycholinguistics*, 34(05), 893-916.

Diskin, C. (2017). The use of the discourse-pragmatic marker ‘like’ by native and non-native speakers of English in Ireland. *Journal of Pragmatics*, 120, 144-157.

Durmuşoğlu Köse, G. and Yüksel, I. (2013). ELT majors' cross sectional evaluation of academic lexical competence and performance. *Journal of Language Teaching and Research*, 4(2), 244.

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.

Erman, B. (1992). Female and male usage of pragmatic expressions in same-sex and mixed-sex interaction. *Language Variation and Change*, 4, 217-234.

Erman, B. (2001). Pragmatic markers revisited with a focus on *you know* in adult and adolescent talk. *Journal of Pragmatics,* 33, 1337-1359.

Fernández-Polo, F. J. (2014). The role of *I mean* in conference presentations by ELF speakers. *English for Specific Purposes*, 34, 58-67.

Fischer, K. and Drescher, M. (1996). Methods for the description of discourse particles: contrastive analysis. *Language Sciences,* 18(3-4), 853-861.

Fischer, K. (2006). Towards an understanding of the spectrum of approaches to discourse particles: Introduction to the volume. In K. Fischer (Ed.), *Approaches to discourse particles* (pp. 1-20). Amsterdam: Elsevier.

Fox Tree, J. E. and 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. *Journal of Pragmatics,* 14, 383-395.

Fraser, B. (1996). Pragmatic markers. *Pragmatics* 6(2), 167-190.

Fraser, B. 1999. What are discourse markers in English?. *Journal of Pragmatics*, 31, 931-952.

Fraser, B. (2010). Pragmatic competence: The case of hedging. In G. Kaltenböck, W. Mihatsch and S. Schneider (Eds.), *New approaches to hedging* (Vol. 9) (pp. 15-34). Bingley: Emerald.

Fuller, J. M. (2003). The influence of speaker roles on discourse marker use. *Journal of Pragmatics,* 35, 23-45.

Fung, L. and Carter, R. (2007). Discourse markers and spoken English: Native and learner use in pedagogic settings. *Applied Linguistics*, 28(3), 410-439.

Gilquin, G. (2016). Discourse markers in L2 English: From classroom to naturalistic input. In O. Timofeeva, A. C. Gardner, A. Honkapohja and S. Chevalier (Eds.), *New approaches to English linguistics: Building bridges* (pp. 213-250). Amsterdam, Philadelphia: John Benjamins.

Götz, S. (2013). *Fluency in native and nonnative English speech* (Vol. 53). Amsterdam, Philadelphia: John Benjamins Publishing.

Halliday, M.A.K. and Hasan, R. (1976). *Cohesion in English*. London: Longman.

Halliday, M.A.K. and Matthiessen, C.M.I.M. (2004). *An Introduction to Functional Grammar*. London, New York: Arnold.

Hansen, M. M.-B. (1998). *The function of discourse particles. A study with special reference to spoken standard French.* Amsterdam, Philadelphia: John Benjamins Publishing Company.

Hays, P. R. (1992). Discourse markers and L2 acquisition. *Papers in Applied Linguistics-Michigan*, 7, 24-34.

Hellermann, J. and 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.

Hernández, T. A. (2008). The effect of explicit instruction and input flood on students' use of Spanish discourse markers on a simulated oral proficiency interview. Hispania, 665-675.

Hernández, T. A. (2011). Re-examining the role of explicit instruction and input flood on the acquisition of Spanish discourse markers. *Language Teaching Research*, 15(2), 159-182.

Hincks, R. (2008). Presenting in English or Swedish: Differences in speaking rate. In A. Eriksson and J. Lindh (Eds), *the Proceedings of Fonetik 2008: The XXIst Swedish Phonetics Conference*, June 11–13 2008, (pp. 21–24). Gothenburg: Reprocentralen, Humanisten, University of Gothenburg.

Huang, L. F. (2011). *Discourse markers in spoken English: A corpus study of native speakers and Chinese non-native speakers.* Unpublished Doctoral Dissertation. Birmingham: University of Birmingham.

Jucker, A. H. 1993. The discourse marker *well*: A relevance-theoretical account. *Journal of Pragmatics,* 19, 435-452.

Jucker, A. H. and Smith, S. W. (1998). "And people just you know like 'wow'": Discourse markers as negotiating strategies. In A. H. Jucker and Y. Ziv (Eds.), *Discourse Markers: Description and Theory* (pp. 171–202). Amsterdam: John Benjamins.

Kim, J. S. (2016). Kim, J. S. (2016). *The effect of instruction on the acquisition of Japanese discourse marker ndesu*. Unpublished Doctoral Dissertation. Urbana: University of Illinois at Urbana-Champaign.

Kormos, J. and Denes, M. (2004). Exploring measures and perceptions of fluency in the speech of second language learners. *System*, *32*(2), 145–164.

Lakoff, R. (1973). Questionable answers and answerable questions. In B. B. Kachru, R. B. Lees, Y. Malkiel, A. Pietrangeli, and S. Saporta (Eds.), *Issues in Linguistics* (pp. 453-467). Papers in Honor of Henry and Renée Kahan. Urbana: University of Illinois Press.

Lam, P. (2006). Well but that's the effect of it: The use of *well* as a discourse particle in talk shows. *International Journal for Language Data Processing*, 30(1), 99-108.

Lam, W. Y. P. (2008). *Discourse particles in an intercultural corpus of spoken English*. Unpublished Doctoral Dissertation. Hong Kong: The Hong Kong Polytechnic University.

Laserna, C. M., Seih, Y. T. and Pennebaker, J. W. (2014). Um... who like says you know: Filler word use as a function of age, gender, and personality. *Journal of Language and Social Psychology*, 33(3), 328-338.

Lee-Wong, S. M. (2001). Coherence, focus and structure: The role of discourse marker ne. *Pragmatics*, 11, 139-154.

Lenk, U. (1998). *Marking discourse coherence. Function of discourse markers in spoken English*. Tübingen: Gunter Narr.

Lennon, P. (2000). The lexical element in spoken second language fluency. In H. Riggenbach (Ed.), *Perspectives on fluency* (pp. 25–42). Ann Arbor, MI: University of Michigan Press.

Levinson, S.C. (1983) *Pragmatics*. Cambridge: Cambridge University Press.

Liao, S. (2009). Variation in the use of discourse markers by Chinese teaching assistants in the US. *Journal of Pragmatics*, 41(7), 1313-1328.

Mauranen, A. (2006). Signaling and preventing misunderstanding in English as lingua franca communication. *International Journal of the Sociology of Language*, 177, 123-150.

Mei, W. (2012). *A Corpus-based comparative study of pragmatic markers: 'I mean' and 'you know' in native and non-native conversation.* Unpublished Doctoral Dissertation. Liverpool: University of Liverpool.

Mullan, K. (2010). *Expressing opinions in French and Australian English discourse: A semantic and interactional analysis*. Amsterdam: John Benjamins.

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 by non-native (Finnish) and native speakers of English. In I. F. Bouton and Y. Kachru (Eds.), *Pragmatics and language learning* (Monograph series No. 4, pp. 126-142). Urbana-Champaign: University of Illinois.

Nikula, T. (1996). *Pragmatic force modifiers: A study in interlanguage pragmatics*. Jyväskylä: University of Jyväskylä.

O'Keeffe, A., McCarthy, M. and Carter, R. (2007). *From corpus to classroom*. Cambridge: Cambridge University Press.

Olynyk, M., d'Anglejan, A. and Sankoff, D. (1987). A quantitative and qualitative analysis of speech markers in the native and second language speech of bilinguals. *Applied psycholinguistics*, *8*(2), 121-136.

Östman, J. O. (1981). *'You know': A discourse-functional study*. Amsterdam: John Benjamins Publishing.

Östman, J.-O. (1982). The symbiotic relationship between pragmatic particles and impromptu speech. In N. Erik Enkvist (Ed.), *Impromptu speech: A symposium* (pp. 147-177). Åbo: The research Institute of the Åbo Akademi Foundation.

Polat, B. (2011). Investigating acquisition of discourse markers through a developmental learner corpus. *Journal of Pragmatics, 43, 3745-3756.*

Redeker, G. (1990). Ideational and pragmatic markers of discourse structure. *Journal of Pragmatics*, 14, 367-381.

Redeker, G. (1991). Review article: Linguistic markers of discourse structure. Review of Deborah Schiffrin, 1987, Discourse markers, Cambridge: CUP. *Linguistics*, 29, 1139-1172.

Redeker, G. (2006). Discourse markers as attentional cues at discourse transitions. In K. Fischer (Ed.), *Approaches to Discourse Particles* (pp. 339-358). Amsterdam: Elsevier.

Romero-Trillo, J. (2002). The pragmatic fossilization of discourse markers in non-native speakers of English. *Journal of Pragmatics*, 34, 769-784.

Rouchota, V. (1996). Discourse connectives: what do they link?. *UCL Working Papers in Linguistics*, 8(199-214), 51.

Sankoff, G., Thibault, P., Nagy, N., Blondeau, H., Fonollosa, M. and Gagnon, L. (1997). Variation in the use of discourse markers in a language contact situation. *Language Variation and Change*, 9(2), 191-217.

Schiffrin, D. (1987). *Discourse markers*. Cambridge: Cambridge University Press.

Schourup, L. (1985). *Common discourse particles in English conversation*. New York: Garland.

Schourup, L. (1999) Discourse markers. *Lingua*, 107, 227-265.

Scott, M. (2011). *WordSmith tools version 6*. Liverpool: Lexical Analysis Software.

Segalowitz, N. (2010). *Cognitive bases of second language fluency*. New York: Routledge.

Shimada, K. (2014). Contrastive interlanguage analysis of discourse markers used by non-native and native English speakers. *JALT Journal*, 36(1), 47.

Simon-Vandenbergen, A.N. (2001). Analysing text and discourse. A report. *The European English Messenger, X*, 79–85.

Skehan, P. (2003). Task-based instruction. *Language Teaching, 36*, 1–14.

Smith, S. W. and Jucker, A. H. (2000). Actually and other markers of an apparent discrepancy between propositional attitudes of conversational partners. In G. Andersen and T. Fretheim (Eds.), *Pragmatic markers and propositional attitude* (pp. 207-238). Amsterdam: John Benjamins.

Svartvik, J. 1980. *Well* in conversation. In Greenbaum, S., G. Leech and J. Svartvik (Eds.), *Studies in English linguistics for Randolph Quirk* (pp. 167-177). London: Longman.

Şener, S. (2015). Vocabulary learning strategy preferences and vocabulary size of pre-service English teachers. *International Journal of Educational Researchers*, 6(3), 15-33.

Taguchi, N. (2002). A comparative analysis of discourse markers in English conversational registers. *Issues in Applied Linguistics*, *13*(1), 41-70.

Tavakoli, P. and Skehan, P. (2005). Strategic planning, task structure, and performance testing. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 239–276). Amsterdam: John Benjamins.

Tekin, M. (2015). Turkish discourse particles 'İşte' and 'falan' and their English Equivalents 'Well' and 'Like': A corpus study. *Tarih Okulu Dergisi*, 435-459.

Tömen, M. (2016). *The relationship between vocabulary size, lexical diversity, lexical density and EFL writing scores: A cross-sectional study*. Unpublished MA Thesis. Eskişehir: Anadolu University, Graduate School of Educational Sciences.

Underhill, R. (1988). Like is, like, focus. *American Speech*, 63, 234–246.

Vandepitte, S. (1993). *A pragmatic study of the expression and interpretation of causality: Conjuncts and conjunctions in modern spoken British English* (Vol. 146). Koninklijke Academie voor Wetenschappen, Letteren en Schone Kunsten van België.

Wei, M. (2011). A comparative study of the oral proficiency of Chinese learners of English across task functions: a discourse marker perspective. *Foreign Language Annals*, 44(4), 674-691.

Xue, W. and Lei, L. (2016). Functions of *you know* and *I mean* in academic seminar discussions by ELF speakers. *Chinese Journal of Applied Linguistics*, 39(3), 288-303.

Zorluel Özer, H. and Okan, Z. (2018). Discourse markers in EFL classrooms: A corpus-driven research. *Journal of Language and Linguistic Studies*, 14(1), 50-66.

APPENDICES

Appendix I. Approval of the Ethics Committee

Kayıt Tarihi: 18.06.2016 | Protokol No: 68885

ANADOLU ÜNİVERSİTESİ ETİK KURULU KARARI

ÇALIŞMANIN TÜRÜ:	BAP Projesi-Doktora Tez Çalışması
KONU:	Eğitim Bilimleri
BAŞLIK:	Discourse Markers In Native And Non-Native Spoken English: A Corpus-Based Comparison Of Turkish And British University Students' Employment Of Discourse Markers In Informal Interviews. (Anadil ve Yabancı Dil Olarak Sözlü İngilizce'de Söylem Belirteçleri: Türk ve İngiliz Üniversite Öğrencilerinin İnformal Görüşmelerde Söylem Belirteçleri Kullanımının Derlem-Temelli Karşılaştırılması)
PROJE/TEZ YÜRÜTÜCÜSÜ:	Prof. Dr. Gül DURMUŞOĞLU KÖSE
TEZ YAZARI:	Yusuf ÖZTÜRK
ALT KOMİSYON GÖRÜŞÜ:	–
KARAR:	Olumlu

ETİK KURUL ÜYELERİ — **İMZA/ TARİH** 30.09.2016

Prof. Dr. Aydın AYBAR
Rektör Yardımcısı / Etik Kurul Başkanı

Prof. Dr. Hayrettin TÜRK
Fen Bil.(Fen Fak.)

Prof. Dr. Yusuf ÖZTÜRK
Sağlık Bil.(Ecz. Fak.)

Prof. Dr. Esra CEYHAN
Eğitim Bil. (Eğitim Bil. Ens.)

Prof. Dr. Bülent GÜNSOY
Sos. Bil.(İkt. Fak.)

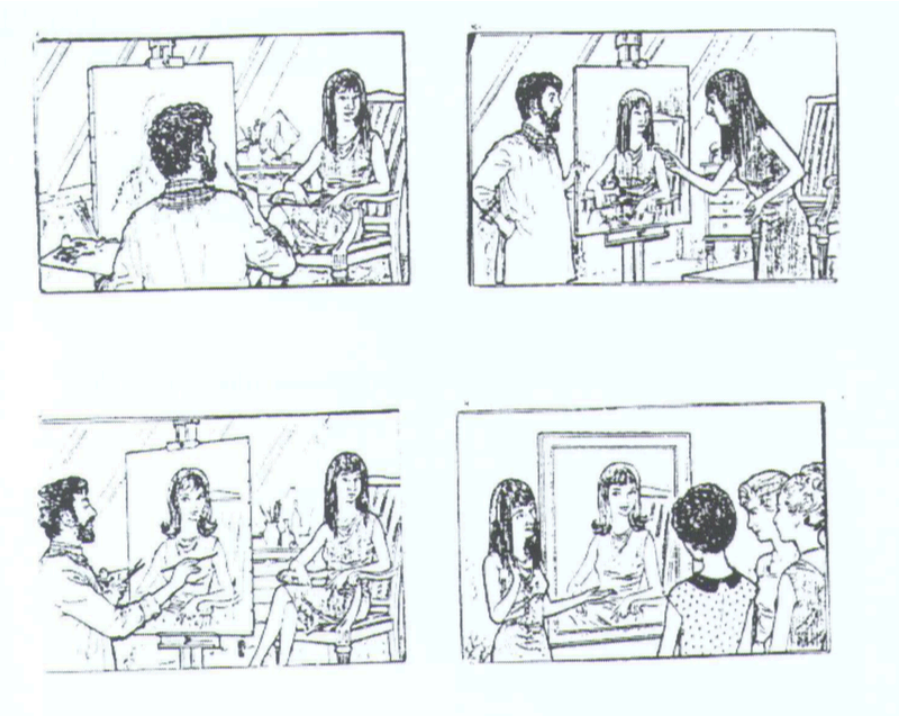
Prof. Dr. Münevver ÇAKI
Güz. San. (Güz. San. Fak.)

Appendix II. Picture-Description Task Used in the Final Part of the Interviews

PICTURE STORY

There are four pictures below that tell a story.

Have a look these pictures and invent a story based on them.



Appendix III. Transcription guidelines

1. Interview identification

Each interview is preceded by a code of this type: <h nt="FR" nr="FR+*three-figure number*">

e.g. <h nt="FR" nr="FR004"> (4th interview with French mother tongue student)

Examples of country codes:

- DUTCH = DU001
- GERMAN = GE001
- NORWEGIAN = NO001
- SPANISH = SP001
- SWEDISH = SW001

All interviews should end with the following tag (on a separate line): </h>

2. Speaker turns

Speaker turns are displayed in vertical format, i.e. one below the other. Whilst the letter "A" enclosed between angle brackets always signifies the interviewer's turn, the letter "B" between angle brackets indicates the interviewee's (learner's) turn. The end of each turn is indicated by either </A> or </B>.

e.g. <A> okay so which topic have you chosen </A>

<B> the film or play that I thought was particularly good or bad really </B>

3. Overlapping speech

The tag <overlap /> (with a space between "overlap" and the slash) is used to indicate the beginning of overlapping speech. It should be indicated in both turns. The end of overlapping speech is not indicated.

e.g. <B> yeah I went on a bus to London once and I'll never <overlap /> do it again </B>

<A> <overlap /> that's even worse </A>

4. Punctuation

No punctuation marks are used to indicate sentence or clause boundaries.

5. Empty pauses

Empty pauses are defined as a blank on the tape, i.e. no sound, or when someone is just breathing.

The following three-tier system is used: one dot for a "short" pause (< 1 second), two dots for a "medium" pause (1-3 seconds) and three dots for "long" pauses (> 3 seconds).

e.g. <B> (erm) .. it’s a British film there aren't many of those these days </B>

6. Filled pauses and backchannelling

Filled pauses and backchannelling are put between brackets and marked as (eh) [brief], (er), (em), (erm), (mm), (uhu) and (mhm). No other fillers should be used.

e.g. <B> yeah . well Namur was warmer (er) it was (eh) a really little town </B>

7. Unclear passages

A three-tier system is used to indicate the length of unclear passages: <X> represents an unclear syllable or sound up to one word, <XX> represents two unclear words, and <XXX> represents more than two words.

e.g. <B> <X> they're just begging <XX> there's there's honestly he did a course .. for a few weeks </B>
If transcribers are not entirely sure of a word or word ending, they should indicate this by having the word directly followed by the symbol <?>.
e.g. <B> I went to see a<?> friend at university there and stayed </B>
Unclear names of towns or titles of films for example may be indicated as <name of city> or <title of film>.
e.g. <B> where else did we go (er) <name of city> it's in Bolivia </B>
8. Anonymisation
Data should be anonymised (names of famous people like singers or actors can be kept). Transcribers can use tags like <first name of interviewee>, <first name and full name of interviewer> or <name of professor> to replace names.
e.g. <A> I'm <first name of interviewer> . what's your name </A>
9. Truncated words
Truncated words are immediately followed by an equals sign.
e.g. <B> it still resem= resembled the theatre </B>
10. Spelling and capitalisation
British spelling conventions should be followed. Capital letters are only kept when required by spelling conventions on certain specific words (proper names, I, Mrs, etc.) – not at the beginning of turns.
11. Contracted forms
All standard contracted forms are retained as they are typical features of speech.
12. Non-standard forms
Non-standard forms that appear in the dictionary are transcribed orthographically in their dictionary accepted way: cos, dunno, gonna, gotta, kinda, wanna and yeah.
13. Acronyms
If acronyms are pronounced as sequences of letters, they are transcribed as a series of upper-case letters separated by spaces.
e.g. <B> yes not really I did sort of basic G C S E French and German </B>

If, on the other hand, acronyms are pronounced as words, they are transcribed as a series of upper-case letters not separated by spaces.
e.g. <A> (mhm) (er) you're doing a MAELT </A>

14. Dates and numbers
Figures have to be written out in words. This avoids the ambiguity of, for example, "1901", which could be spoken in a number of different ways.
e.g. <B> an awful lot of people complain and say well the grants were two thousand two hundred </B>
15. Foreign words and pronunciation
Foreign words are indicated by <foreign> (before the word) and </foreign> (after the word).
e.g. <B> we couldn't go with (er) knives and so on <foreign> enfin </foreign> we were (er) </B>
As a rule, foreign pronunciation is not noted, except in the case where the foreign word

and the English word are identical. If in this case the word is pronounced as a foreign word, this is also marked using the <foreign> tag.
e.g. <B> I didn't have the (erm) . <foreign> distinction </foreign> </B>
16. Phonetic features
(a) Syllable lengthening
A colon is added at the end of a word to indicate that the last syllable is lengthened. It is typically used with small words like to, so or or. Colons should not be inserted within words.
e.g. <B> that's something I'll I'll plan to: to learn </B>
(b) Articles
- when pronounced as [ei], the article *a* is transcribed as a[ei];
e.g. <B> and it's about (erm) . life in a[ei] (eh) public school in America I think </B>
- when pronounced as [i:], the article *the* is transcribed as the[i:].
e.g. <B> and the[i:] villa we were staying in was in one of the valleys </B>
17. Prosodic information: voice quality
If a particular stretch of text is said laughing or whispering for instance, this is marked by inserting <starts laughing> or <starts whispering> immediately before the specific stretch of speech and <stops laughing> or <stops whispering> at the end of it.
e.g. <B> <starts laughing> I don't have to assess it I only have to write it <stops laughing> </B>
18. Nonverbal vocal sounds
Nonverbal vocal sounds are enclosed between angle brackets.
e.g. <B> I hope so I've I've got some <coughs> friends out there </B>
e.g. <B> so I went back into Breda .. and sat down again <imitates the sound of a guitar> </B>
19. Contextual comments
Non-linguistic events are indicated between angle brackets only if they are deemed relevant to the interaction (if one of the participants reacts to it, for example).
e.g. <A> no it's true it's nice to have your own bathroom </A>
<somebody enters the room>
<B> hi </B>
20. Tasks
The three tasks making up the interview (set topic, free discussion and picture description) should be separated from each other. This is done using the following tags: <S> (before the set topic), </S> (after the set topic), <F> (before the free discussion), </F> (after the free discussion), <P> (before the picture description), </P> (after the picture description). These tags should occupy a separate line and should not interrupt a turn.
e.g. <S>
<A> did you . manage to choose a topic </A>

Appendix IV. Sample Transcription of an Interview in the Turkish Corpus

TR Participant 1

00:00:21.19 <A> okay thank you again (er) I'm with <first name of interviewee> and today is the November twenty-fifth thank you again for your participation (er) this is not a speaking test okay </A> 00:00:34.14
00:00:34.14 <B> okay </B> 00:00:34.14
00:00:33.15 <A> we're just having an informal chat (er) I will show you three topics (er) you will read them and you will think about then you can choose one and start talking about it and then we will have a conversation about it okay </A> 00:00:53.12
00:00:53.12 <B> okay I get it I guess </B> 00:00:54.15
00:00:54.21 <A> <laughing> </A> 00:00:56.05
00:01:00.10 <B> an experience .. well I can think about . a film that impressed me . (er) so should I think for a moment <overlap/> or just </B> 00:01:16.08
00:01:15.15 <A> <overlap/> yeah you can you can think for a moment </A> 00:01:17.16
00:01:19.14 <B> well I can go with the lobster have you watched it before </B> 00:01:22.27
00:01:22.27 <A> it sounds familier yeah </A> 00:01:26.00
00:01:26.00 <B> (er) there was a hotel and they . kindda took people who didn't have maids . didn't have wives husbounds or </B> 00:01:38.20
00:01:38.20 <A> yeah yeah </A> 00:01:39.20
00:01:39.20 <B> you know that </B> 00:01:40.11
00:01:40.11 <A> I saw it like three or four weeks ago <overlap/> so I remember yeah </A> 00:01:48.26
00:01:44.23 <B> <overlap/> three or four weeks wow that's very </B> 00:01:47.27
00:01:50.03 <B> (erm) it was really good because I think they captured the reality in a way that had black homour you know and I'm really interested in that so I liked it and the way they showed the rebellion reb=rebels in that movie and how they treated in the same way (er) . like they did in the hotel they are doing the same thing but in a different </B> 00:02:29.00
00:02:29.26 <A> maybe the opposite way </A> 00:02:32.02
00:02:32.02 <B> yes <laughing> but you know the treatment was the same actually so: . <X> standards . they show that very well and I liked it what else . I can say about that </B> 00:02:46.25
00:02:46.25 <A> which side would want to be if you were in that <laughing> </A> 00:02:51.17
00:02:51.17 <B> well definetely not the hotel <laughing> (erm) . not the woods too because you know you can't interact with the opposite sex and their limitations . and their kind of . observing you like what are you gonna do with that boy <laughing> what is the relationship between you </B> 00:03:20.09
00:03:21.05 <A> perhaps it makes it more exciting maybe right <laughing> </A> 00:03:23.25
00:03:23.25 <B> well maybe but you know it was kind of ridiculous because (er) the man in the movie was . had a sight problem . and you know the other woman who also had the same problem was attracted to him because they had the same health issue and I think that represents how we choose our . maids <laughing> it's kind of ridiculous </B> 00:03:56.04
00:03:56.04 <A> did they have the sight problem from the beginning or </A> 00:03:59.11
00:04:00.08 <B> yes they were both <overlap/> farsigheted </B> 00:04:04.21
00:04:01.26 <A> I think there was some sort of punishment right </A> 00:04:06.24
00:04:07.07 <B> before that the woman also had the same problem but she didn't use glasses she just couldn't see very well . and the man had glasses so they were kind of they connected over that . I remember like that but maybe . and after that after the punishment woman became blind and she wanted the same thing she wanted him to be blind to so they could go on <starts laughing> the relationship <stops laughing> but (er) it was open ended we didn't know if he did . he was using a knife and he was going to stab his own eyes </B> 00:04:55.15
00:04:56.23 <A> were they looking for some sort of treatment for it </A> 00:04:59.19
00:05:01.16 <B> not really he had a knife and he went to the bathroom and he was looking at the mirror and thinking if he should you know stab himself and not and the movie finished there . it's how I remember I don't know </B> 00:05:19.00
00:05:22.03 <A> well . I think so . yeah I don't remember well because I mean . I don't know where I wathced it or whom I watched it with who did you wathch it with </A> 00:05:35.19
00:05:35.19 <B> I was alone <starts laughing> <X> <stops laughing> </B> 00:05:37.25
00:05:37.25 <A> how did you manage to choose such a movie </A> 00:05:40.07
00:05:40.07 <B> well I like those <starts laughing> ind of movies <stops laughing> (erm) </B> 00:05:42.15
00:05:42.15 <A> there is a name for that kind of movies </A> 00:05:45.27
00:05:46.05 <B> festival movies I don't </B> 00:05:47.25
00:05:47.25 <A> no not festival movies I don't remember now </A> 00:05:50.11
00:05:50.11 <B> distopian <overlap/> distopian </B> 00:05:52.01
00:05:50.22 <A> <overlap/> ha distopia yeah distopia </A> 00:05:51.25
00:05:50.22 <A> not utopia but distopia </A> 00:05:56.06
00:05:56.06 <B> distopia </B> 00:05:55.27
00:05:55.27 <A> I like that kind of movies </A> 00:05:58.18
00:05:56.29 <B> yeah me too it's interesting to see how world works in a different way <laughing> like in a hotel in the woods </B> 00:06:05.17
00:06:05.17 <A> you someitmes imagine yourself I mean what would I do if I were in that situation right </A> 00:06:11.05
00:06:11.05 <B> yeah what would I be if I were an animal . I probably be a dog </B> 00:06:19.03
00:06:18.02 <A> they turn people into an animanl right </A> 00:06:19.21
00:06:21.07 <B> yes yes and he want to become a lobster . that's . <starts laughing> interesting choice <stops laughing> but I don't know I wouldn't be a lobster </B> 00:06:29.19
00:06:30.06 <A> what would you be </A> 00:06:30.21
00:06:31.17 <B> I would be a dog I guess I love dogs <overlap/> what would you be </B> 00:06:35.27
00:06:35.07 <A> <overlap/> have your ever </A> 00:06:36.02
00:06:35.04 <B> what would you be </B> 00:06:35.29
00:06:36.05 <A> have you ever had a dog </A> 00:06:38.00

00:06:39.04 <B> I had in back at Izmir but it was my uncle's dog and every summer we . we visit them and I can you know play with him . <starts laughing> he grows up super fast <stops laughing> he was like four months old and he was this big and then I saw him after I don't know six months or something and he was like this <starts laughing> what did you feed him <stops laughing> I was like that how is it possible </B> 00:07:16.23
00:07:16.23 <A> perhaps it's the kind of the dog or maybe every dog grows up like that I don't know </A> 00:07:22.19
00:07:23.04 <B> I don't know I was really shocked he was this big and then I saw him and he was like what how did that happen did you change the dogs but yeah he grows up really fast and he doesn't know that he grows up so he thinks that he is still this big and he kindda attacks <overlap/> <starts laughing> jumps on people <stops laughing> and he wants to play and we're like stop stop <starts laughing> you're hurting me hurt you're hurting me <stops laughing> please stop and he just . he's like ha everything is a game </B> 00:07:56.16
00:07:57.12 <A> how old is he now </A> 00:07:58.14
00:08:00.02 <B> (erm) I guess he is one . he is one years old </B> 00:08:05.03
00:08:05.03 <A> okay </A> 00:08:06.22
00:08:06.18 <B> his first year on earth </B> 00:08:08.21
00:08:12.03 <A> (er) do you remember any other movies like that </A> 00:08:15.04
00:08:15.04 <B> any other movies like . like the lobster or </B> 00:08:18.07
00:08:18.07 <A> I'm surprised that you mentioned lobster as your favorite movie . or a movie that impressed you </A> 00:08:24.17
00:08:24.17 <B> well I guess it's because (er) recently I haven't seen a movie that affects me that much and I watched it this summer I guess . so . it just came to my mind </B> 00:08:42.11
00:08:43.10 <A> what about tv series did you follow any </A> 00:08:45.17
00:08:47.00 <B> I used to I used to watch stupid tv shows but then I just dropped it </B> 00:08:54.01
00:08:54.25 <A> why stupid </A> 00:08:56.02
00:08:56.02 <B> well because you know there are some tv shows that . just goes on like eight seasons or ten seasons and it just become boring really and it just doesn't go anywhere like a circle <starts laughing> stupid circles <stops laughing> just repeats itself </B> 00:09:18.22
00:09:19.07 <A> really </A> 00:09:19.28
00:09:21.08 <B> <laughing> I just I got bored I guess so . I'm more into movies . now . but I used to watch game of thrones </B> 00:09:32.07
00:09:33.10 <A> you stopped </A> 00:09:34.08
00:09:34.25 <B> yes <overlap/> everyone is shocked </B> 00:09:36.27
00:09:35.28 <A> <overlap/> wow </A> 00:09:35.28
00:09:37.02 <A> it must be hard </A> 00:09:37.20
00:09:37.20 <B> <laughing> how did you stop . yeah I stopped I stopped after season two </B> 00:09:45.04
00:09:45.24 <A> season two </A> 00:09:46.10
00:09:46.10 <B> season two yes </B> 00:09:48.12
00:09:48.12 <A> wow that's where it all starts . after season three I guess . </A> 00:09:53.10
00:09:54.15 <B> and it's on . its sixth season </B> 00:09:58.25
00:09:58.25 <A> I think <overlap/> I'm not sure I also watch walking dead so . I'm not sure </A> 00:10:05.07
00:10:05.07 <B> <overlap/> or fifth </B> 00:10:00.21
00:10:05.11 <B> walking dead I I just hate that <starts laughing> tv show <stops laughing> it's just . kindda disgusting and you know . does it have a story </B> 00:10:16.18
00:10:16.18 <A> it does actually I there is a certain group of have you ever watched it </A> 00:10:23.04
00:10:23.25 <B> just a couple of episodes </B> 00:10:26.07
00:10:26.07 <A> there is a certain group of people and they just happen they just try to survive you know in a world like that I mean it's also something like distophia because . I mean it's not really realistic but you think that I mean what would I do if I had to be in that situation because it's not really dead people they fear of it's the people who are alive that they fear of because there is no authority no police no soldiers no army anything and people become . like barbarians I mean they just (er) rob <overlap/> each other </A> 00:11:08.02
00:11:06.17 <B> <overlap/> they create their own justice </B> 00:11:08.10
00:11:08.28 <A> yeah yeah exactly . they rob each other kill each other for their food I mean no one you can't trust no one it's just an awful situation . so (er) </A> 00:11:22.03
00:11:22.03 <B> and you like watching that <laughing> </B> 00:11:24.13
00:11:24.13 <A> I do actually </A> 00:11:24.28
00:11:25.16 <B> wow </B> 00:11:25.27
00:11:26.13 <A> I mean it's it's exciting to see how people try to survive so . I don't know <laughing> </A> 00:11:32.15
00:11:33.22 <B> well everyone likes different things I guess (erm) </B> 00:11:38.21
00:11:39.03 <A> so you never follow any foreign tv series </A> 00:11:42.10
00:11:42.27 <B> I used to friends how I met your mother (erm) </B> 00:11:46.19
00:11:46.19 <A> those are like sitcoms right </A> 00:11:48.10
00:11:48.10 <B> yes sitcoms </B> 00:11:49.17
00:11:50.28 <A> have you finished how I met your mother </A> 00:11:53.04
00:11:53.04 <B> no but I finished friends I after I watched friends I hated how I met your mother <overlap/> <starts laughing> I don't know why <stops laughing> </B> 00:12:01.29
00:12:00.05 <A> <overlap/> really </A> 00:12:00.12
00:12:01.15 <B> I guess it's . because it's not original . </B> 00:12:06.00
00:12:07.08 <A> like a copy cat of friends </A> 00:12:08.23
00:12:09.11 <B> yes yes I think that way . I don't know . I don't like how I met your mother . and well what else lucifer you know lucifer </B> 00:12:23.11
00:12:23.11 <A> well I keep hearing about it but I never watched it actually . are you remembering the tv series called super natural </A> 00:12:31.02
00:12:31.02 <B> yes that's </B> 00:12:32.11
00:12:32.11 <A> is it similar to lucifer . or </A> 00:12:34.16
00:12:35.03 <B> no not really</B> 00:12:36.07
00:12:36.07 <A> because lucifer was kept mentioning in supernatural . in different ways </A> 00:12:40.08
00:12:40.18 <B> well (er) in supernatural it wasn't the main theme but after season three or something it became the main </B> 00:12:52.13

00:12:52.13 <A> lucifer <laughing> </A> 00:12:53.24
00:12:54.01 <B> yes . no in supernatural . the angels and deamons they just came pop into <starts laughing> <X> <stops laughing> they were like after season three all of the show was about angels and deamons but in lucifer there is a guy and he is he is the real lucifer he is an angel or the devil I don't know they kindda represent him as the deamon I guess the devil and it just starts like . lucifer gets bored and he just comes to earth and <starts laughing> he does something <stops laughing> like I don't know he had a friend back in . hell </B> 00:13:43.17
00:13:44.03 <A> yeah </A> 00:13:44.11
00:13:44.26 <B> and they kindda . have quarrels . and he starts <overlap/> seeing the </B> 00:13:53.17
00:13:52.11 <A> <overlap/> in the world </A> 00:13:52.25
00:13:54.09 <B> yes . in the earth </B> 00:13:55.27
00:13:56.03 <A> they come to world come to earth </A> 00:13:57.04
00:13:57.04 <B> yes . it's . kindda criminal show because he meets a police . a woman and he is kindda attracted to her and he . he just . I don't know it's wierd </B> 00:14:16.13
00:14:16.13 <A> sounds interesting I think I should give it a try </A> 00:14:19.20
00:14:20.23 <B> he is the devil and he is in he is on earh and </B> 00:14:24.04
00:14:25.01 <A> he is also attracted to a woman </A> 00:14:26.14
00:14:26.14 <B> he is attracted to a human . human being and he can't read her so: that confuses him and </B> 00:14:36.06
00:14:36.06 <A> like read her mind </A> 00:14:37.17
00:14:38.00 <B> yes he can change people's ideas he can control <overlap/> their minds </B> 00:14:42.16
00:14:41.23 <A> affect them </A> 00:14:42.07
00:14:42.27 <B> yes affect them but he can't do that with that woman and . I think he is interested in her so he stays with her so: . he helps her . with her cases with the NYPD or something </B> 00:15:03.14
00:15:03.14 <A> yeah new york police department </A> 00:15:05.05
00:15:05.05 <B> yes they were in new york or los angles I can't remember . maybe LA </B> 00:15:10.27
00:15:10.19 <A> LAPD </A> 00:15:11.26
00:15:11.26 <B> I don't know <laughing> something like that so it's a bit criminal and it's a bit supernatural . and I hate supernatural the tv show <laughing> </B> 00:15:20.02
00:15:20.18 <A> you hated it </A> 00:15:21.02
00:15:21.02 <B> yes . just what I mentioned you know the circle and it <starts laughing> just doesn't go anywhere <stops laughing> and I don't like it </B> 00:15:30.04
00:15:30.12 <A> yeah I mean I watched supernatural for like three four seasons and I realised that circle and then I just stopped where is it going nowhere . is it finished . supernatural </A> 00:15:42.04
00:15:42.04 <B> no it's still going on and it's on its twelveth season . twelveth yes </B> 00:15:50.26
00:15:50.26 <A> it must be </A> 00:15:51.18
00:15:53.12 <B> and I realised that <laughing> eigth season so I was really late . you were early that's good </B> 00:16:02.09
00:16:02.10 <A> yeah (er) there is one last thing .. (er) I'll show you four pictures </A> 00:16:09.11
00:16:09.21 <B> okay </B> 00:16:10.04
00:16:10.04 <A> okay . you can just examine them . as long as you want . and then make up a story out of it </A> 00:16:20.19
00:16:20.19 <B> make up a story of it </B> 00:16:22.07
00:16:22.07 <A> it doesn't have to be long .. create your own circle <laughing> </A> 00:16:30.06
00:16:30.06 <B> that's a good one <laughing> . okay </B> 00:16:31.24
00:16:43.03 <A> I'm sorry it's not very clear but I think you can get </A> 00:16:45.21
00:16:45.21 <B> yeah I .. once upon a time there was a selfish . concited woman who . want to . </B> 00:16:59.27
00:17:02.08 <A> it doesn't have to be like a past time story </A> 00:17:04.01
00:17:04.28 <B> wants to . want a picture of herself so: she goes to a painter and she asks him to paint herself . and </B> 00:17:21.17
00:17:21.17 <A> like a portrait </A> 00:17:22.17
00:17:22.17 <B> yes like a portrait so: . he says okay I'll do that <laughing> and he draws her . but she doesn't like it so: she wants to take the picture to her friends . and she asks them if it's . if it represents the reality . like am I really like this am I this ugly . no I'm the . most perfect girl in the world <laughing> so . and her friends examine the picture </B> 00:18:06.23
00:18:06.23 <A> okay </A> 00:18:07.12
00:18:07.12 <B> and .. they think it's amazing because she's really ugly and <X> and they have a fight <laughing> .. I don't know . <starts laughing> I'm really making up right now . nothing else come <stops laughing> </B> 00:18:28.12
00:18:28.24 <A> that sounds like a circle <laughing> </A> 00:18:30.25
00:18:30.25 <B> well .. so they have . yeah that's all I don't know <laughing> </B> 00:18:39.17
00:18:39.17 <A> okay thank you </A> 00:18:43.16

Appendix V. Sample Transcription of an Interview in the LOCNEC

NS Participant 1

<A> at which distance we have to: . right ... erm what are you going to talk to me about <\A><B> I thought I'd talk about a f= a film which I have seen recently <\B><A> uhu <\A><B> which is erm . Dead Poets' Society <\B><A> oh yes <\A><B> it's a film which I . I had never seen before but everyone else seemed to have seen <laughs> and so I was interested to see it . and it's about erm . life in a[ei] . eh public school in America I think it's in about the nineteen fifties <\B><A> yeah I think so yeah <\A><B> erm and .. it . it shows how erm . a a group of boys inspired by a sort of . unauth= . unauthentic sort of English t not unauthentic but erm . <\B><A> [mhm <\A><B> [unorthodox English teacher . erm form a group called Dead Poets Society <\B><A> mhm <\A><B> where they erm .. read poetry and and treat it as a . a living thing . and erm .. one of the boys is particularly inspired by this erm . besides <X> he wanted to become an actor <\B><A> mhm <\A><B> erm his parents are very much against this because they've had lots of . plans . for his life how he's going to become a doctor <\B><A> mhm <\A><B> erm because he's <\B><A> just like his father <\A><B> yes and he and well he's he's having chances that his parents never had his father had . worked from being very . very poor and humble and . and managed to drag himself up and erm .. and . they think they've giving their son all these chances they never had <\B><A> mhm <B> and so they're very keen that erm . they should do .. or that<?> he should do what they want him to <\B><A> mhm <\A><B> erm but he decides he wants to become an actor . and erm he takes part in a play against his erm . father's wishes and er . his father's furious with him and eventually the boy commits suicide <\B><A> [yeah <\A><B> [and it's very sad at <begin_laughter> the[i:] end <\B><A> yeah it's very sad isn't <laughs> <\A><B> <X> <end_laughter> but erm I thought it was very interesting because it erm .. highlighted the sort of gap that can exist between . generations <\B><A> mhm <\A><B> and and how erm .. sometimes parents want to fulfil all their wishes <\B><A> mhm <\A><B> [<X> <\B><A> [through their . children <\A><B> yes and the things they never did <\B><A> uhu <\A><B> through their children and erm .. I just thought it was very interesting because of that it was erm .. it was<?> quite a it was<?> quite a sad and touching <begin_laughter> <XX> <end_laughter> <\B><A> yeah it is it is I was disappointed at the[i:] end <X> I thought oh it's great and then <X> commits suicide [oh no <X> everyone goes like <X> what's going on now <\A><B> [<laughs> <\B><A> but you were talking about generation gap . did you: do you have . the same kind of experience ... with your parents or grandparents <\A><B> er I I've never really had that that problem .. myself erm . I don't think I think I get on . well with my parents <\B><A> mhm <\A><B> and erm . and they're quite accepting of you know whether I wanted to er go to university or not go to university cos . for a while I didn't want to . go to <\B><A> [mhm <\A><B> [university at all and my parents were all right about that they just kind of said you know well what would you like to do .. erm so it's not a problem I've ever had but I have seen it in . you know some of my friends <\B><A> uhu <\A><B> where erm .. you know they've been . they've felt forced into going to university by their parents <\B><A> even though they didn't want really to go <\A><B> and then maybe other people erm .. have wanted to go to university but because their parents thought it was . <begin_laughter> useless and pointless <end_laughter> <\B><A> yeah . uhu <\A><B> they haven't felt supported <\B><A> so what made you: change your mind and come to university afterwards <\A><B> <laughs> erm . there was a number of things really erm ... one was this er .. quite a practical reason really this erm . the job market was <begin_laughter> so very <end_laughter> bad when I left school <\B><A> uhu <\A><B> erm . <X> I I didn't think I would have much chance of . er getting a job without any further qualifications . and also I thought . I would like . I would like to do more learning than I had . done <\B><A> uhu <\A><B> at school erm . <X> I quite en= enjoy it . in some <begin_laughter> ways <end_laughter> it <X> <\B><A> mhm . and er why did you choose linguistic I think you're a linguistics student <\A><B> erm . I I'm doing a . linguistics minor . erm as part of .. er <\B><A> and what are you doing <\A><B> oh actually it's . I don't know if it counts as a minor itself it's part of English literature . [erm <\B><A>
[ah so you're doing literature and you're doing some courses in linguistics <\A><B> yes [yeah <\B><A> [m h m <\A><B> er just the one in fact er <\B><A> just okay <\A><B> yeah <\B><A> uhu .. and er why did you choose literature <\A><B> .. erm .. well <\B><A> <begin_laughter> good question <end_laughter> <\A><B> I I've always been . erm . very keen on reading <\B><A> [mhm <\A><B> [and<?> .. and in my first year I did .. English literature and language . and French so there was reading involved in . most of my courses really <\B><A> so it's your second year here <\A><B> it's my second year here <\B><A> oh yeah mhm <\A><B> yeah .. erm .. and I found it hard to decide which . course to carry on with cos I didn't really want to . drop French <begin_laughter> but er <end_laughter> <\B><A> uhu <\A><B> erm .. but in the[i:] end I decided that er<?> if I was going to do French I'd probably have wanted to do it . for a as a complete <\B><A> [uhu <\A><B> [erm degree scheme <\B><A> so you dropped [it <\A><B> [and .. I I dropped it rather than rather than do it as a minor <\B><A> mhm <\A><B> because erm I would have . if I'd done it . along with English cos I didn't want to drop English . erm . I probably would have ended not not doing the[i:] year abroad for the French and so maybe I'd have got behind <\B><A> mhm <\A><B> . the[i:] other people <\B><A> what do you wanna do er when you . you leave university .. what kind of job would you like to have or would you like to carry on studying or perhaps go a year abroad or <\A><B> erm . I don't think I want to carry on . studying erm . an academic subject as such but erm .. I'm quite interested in doing something like occupational therapy . and that would involve doing a . another course at . er university <\B><A> w= what is it because I don't know eh . what occupational therapy is <\A><B> . it it's to do with the[i:] idea that erm people's erm mental and physical health can be improved erm . with suitable activity like people who have suffered stokes and that sort of thing <\B><A> uhu <\A><B> er who have erm been ill or have some sort

of disability and erm . helping them to learn to do things that they used to do before they were ill <\B><A> oh yeah <\A><B> maybe <\B><A> that's interesting <\A><B> writing erm <\B><A> [mhm <\A><B> [or or whatever <\B><A> mhm <\A><B> and that sort of thing <\B><A> and do you have to: to study for one more year to do that perhaps <\A><B> erm .. if you do it as a first degree it's . three years but I think it's . two years if you do it after .. [you've already done a degree <\B><A> [uhu .. uhu so you'd like to do that <\A><B> yes I'd like to <\B><A> uhu . erm .. so are you staying on campus here <\A><B> er n= no I er <\B><A> no <\A><B> I live in town <\B><A> why <\A><B> . er well being a second year there are= there aren't any places on campus [for <XX> <\B><A> [so you you're not allowed to have a room on campus . during your second year <\A><B> no . no <\B><A> oh so oh I didn't know that <\A><B> oh right <laughs> <\B><A> oh and erm .. is it very different from living on campus when you live in town what do you prefer <\A><B> erm <\B><A> . if you have any preferences <laughs> <\A><B> it it is it is quite nice living off campus because erm .. you don't feel erm . stuck on campus all the time it can get a little bit claustrophobic I found . last year <\B><A> mhm <\A><B> and you you go home to a different place at the[i:] end of the day . but there are disadvantages <begin_laughter> like it's more <end_laughter> expensive <\B><A> yeah you have to take the bus .. I guess <\A><B> erm . yeah . erm and er ... generally rent works out <?> more expensive <\B><A> [yes <\A><B> [and the bills and everything<?> <\B><A> mhm <\A><B> but erm .. no it is nice erm .. living . living in a house <\B><A> in a house it's in a house <\A><B> yeah <\B><A> with <X> other students or just with a family or .. with other people who: who do something else <\A><B> .. who who do I live with <\B><A> yeah <\A><B> erm I live with erm .. two other students and . the[i:] fiancee of one of those [students <\B><A> [oh yeah [<laughs> <\A><B> [<laughs> <\B><A> did you know them before you: you moved into the house <\A><B> er yes <\B><A> oh yes so you <\A><B> . yeah erm one of . one of them was already living in the house with some other friends who are now living elsewhere so . I knew from them what the house was like so <\B><A> uhu and er do you the city of Lancaster .. the town <\A><B> . yeah yes I do like it erm ... it .. it's nice to go into the town from campus sometimes because the campus is so modern . and in . in the town [there's a lot of older buildings [which is nice<?> <laughs> <\B><A> [yes .. yeah with the castle and that kind of thing yeah <\A><B> and the and the shops are quite nice and <\B><A> yeah it's better than here on campus <\A><B> mhm <laughs> <\B><A> .. because here if you want to: to go out at weekends or if you want to: to go out eating you can't everything is closed you . you have to cook for yourself or just go into town it's a bit . sometimes you know <\A><B> yes <\B><A> erm I think you're a member of the Film Society if you went to see <\A><B> [yes <\B><A> [Dead Poets' Society <\A><B> yes <\B><A> erm . are you a member of .. another society <\A><B> er yeah I'm I'm in the Choral Society <\B><A> the Choral .. so singing <\A><B> yes <\B><A> and what does that involve <\A><B> erm ... well we have a we have a concert about . every . every term . we <X> we've just done one and we did erm .. some Haendel and also a piece by . Vaughan Williams and it's it's mostly that sort of more serious sort of music <\B><A> mhm <\A><B> erm but<?> I really enjoy that eh <\B><A> when was it <\A><B> . it . it was erm .. Sunday first<?> <\B><A> mm ... I didn't see anything for that<?> [<laughs> you should advertised it more <laughs> <\A><B> [<laughs> <\B><A> and probably we . more people would come I don't know [<X> <\A><B> [yeah <\B><A> if a lot of people came to: to see it <\A><B> no there weren't there weren't a very lot of <begin_laughter> people actually no <end_laughter> <\B><A> where was it in . [Great Hall <\A><B> [it was in Great Hall yes <\B><A> mhm .. and you enjoy singing <\A><B> yeah [yes I do <\B><A> [yeah <\A><B> I used to have erm . singing lessons when I was at school <\B><A> [oh yeah <\A><B> [but er I can't really afford to <begin_laughter> now <end_laughter> but er . but I'd like to again some time <\B><A> so you must be a pretty good singer then <\A><B> well no I'm<?> not quite good but I enjoy it <laughs> <\B><A> well I gue= I guess you're very modest <laughs> but do you find it easy to combine all those activities with your academic work <\A><B> . erm .. I suppose sometimes there are . problems fitting everything in . timewise but erm . but having some hobbies . as I do it it sometimes makes it easier to work if I you know spend an evening doing something else <\B><A> mhm <\A><B> you know you can forget about work for a while and <\B><A> get your mind off things [<laughs> <\A><B> [yes and so then it's easier to come back to work afterwards <\B><A> <X> you feel more relaxed and . yeah <XX> <laughs> I think you come from M= Manchester well not too far from Manchester do you go back every weekend . or are you staying on well in town <laughs> <\A><B> <laughs> erm no I have I have been home once this term but er . I don't normally go home very much erm . well I go home for the hol= holidays erm <\B><A> so Christmas Easter <\A><B> yes <X> <\B><A> summer holidays <\A><B> yes <X> <\B><A> and that's it <\A><B> .. yeah well the the holidays are are quite .. long <X> we have a month at Christmas a month at Easter <\B><A> mm <\A><B> and then <X> three months in the summer <X> and <\B><A> mhm <\A><B> my my parents tend to come up once or twice a term <\B><A> oh yes <\A><B> to: see me cos it's not all that far to come <\B><A> no no it's <X> I mean one and a half hours by train or something like that <\A><B> yeah yeah <\B><A> so I mean it should be doable <laughs> because well in Belgium we go back .. well say every weekend <\A><B> mm <\B><A> . even i= well . especially if you live only: one hour and a half away from the[i:] university so .. is it <\A><B> yes <\B><A> .. usual here to: to stay on campus or in town or whatever during the weekend <\A><B> erm . yeah I think quite quite a lot of people . erm stay here for most of the term <\B><A> mhm <\A><B> erm <\B><A> .. it's funny <laughs> <\A><B> <laughs> <\B><A> I mean it's as if you were living here really <\A><B> mm <\B><A> because well when we: .. well in Belgium when we go back . we we sort of . we have two homes in a way <\A><B> mhm <\B><A> and we we . we still have a lot of activities in .. our home towns <\A><B> mm <\B><A> and a few on campus . so here I think you have all your activities here <\A><B> mm <\B><A> in Lancaster and then forget [everything about Manchester or <\A><B> [yes ... yeah it it is quite odd when you erm . spend the whole term . here and then . go home and you're there for a month it's as if you have two sort of

separate . worlds in a way <\B><A> uhu <\A><B> and then at first I found that quite difficult when I was first at university but now that erm I know more people and they come to visit me at home <\B><A> [mhm <\A><B> [and I go to their homes on holidays . it you know it's as if the two worlds mix a lot more <\B><A> mhm cos I mean .. it must be difficult living on campus for two months and then going back <\A><B> mhm <\B><A> home I mean you're independent here you can do whatever you want to and then [you go back home <\A><B>
[yes .. mhm <\B><A> how do you feel about that is it sometimes difficult I mean . you have to to I guess to tell your parents where you're going to if you leave and that kind of thing <\A><B> .. erm ... yeah it it is it is quite . difficult to I suppose it's something I've got used to a lot more I do I do like going home it has it has advan= some advantages over being here and being here <\B><A> you don't have to cook <laughs> <\A><B> <begin_laughter> well I do have to do some cooking <end_laughter> but <\B><A> yeah I mean but <\A><B> yeah not so much yeah [so <\B><A> [not so much <\A><B> er .. yeah I I like going home <X> I do get on with my parents and they're not they're not very . strict but erm yes I d= I do . feel yeah I do have to . tell them . where I'm going and <\B><A> yeah [<X> <\A><B> [and that sort of thing so it is a bit different but erm <\B><A> you have get used to it <\A><B> mhm <\B><A> mhm <\A><B> yeah <\B><A> er do you have brothers and sisters <\A><B> I have got . one sister yes er [she's . she's older than me yes <\B><A> [bigger ... she's [older <\A><B>
[she's .. erm she's doing an M A at the moment in Manchester so she: she doesn't live with my parents but she's there quite a lot of the time <\B><A> uhu <\A><B> erm .. and she's doing <\B><A> do you get her very well or <\A><B> yes yes I get on with her [well<?> <\B><A> [do you miss her when you you're not here and she's not there in .. well <\A><B> [yes <\B><A> [<X> <laughs> <\A><B> <laughs> yes I I do miss her yes but erm . erm . I suppose it's been this situation for quite a while because .. last year I was here while she was living well she was living at home and working and then before that she was doing her first degree <\B><A> mhm <\A><B> while I was still at school so for the past five years we've not really be living in the same place but erm . yeah it's nice to see her when I do see her <\B><A> mhm I can imagine <\A><B> mm <\B><A> well . before you go I want you to have a look at this . it's a story so you have four pictures <\A><B> mhm <\B><A> and I'd like you to tell me that story <\A><B> ... erm ... the man is drawing a . a picture of maybe it's his wife or his <\B><A> [mhm <laughs> <\A><B> [his girlfriend and erm and she's she's posing for him to do the picture and then she sees what he's drawn and .. she doesn't like it and she's she's cross with him and .. it it does look quite like her <\B><A> [<laughs> <\A><B> [but she thinks it's .. it's too ugly erm .. so then it looks like he's drawn the picture again but erm it draws it making her look more beautiful and she's<?> he's made a . smile where<?> she's really sort of frowning <\B><A> [<laughs> <\A><B> [<X> made her hair curl <\B><A> mhm <\A><B> erm .. and then she s= she seems pleased with that picture . and she's showing it off to all her friends so it's obvious that the picture looks quite different from her <\B> <A> mhm . okay thank you very much <\A>