A study on metadiscoursive interaction in the doctoral dissertations of the native speakers of English and the Turkish speakers of English

Zehra Köroğlu [a] *

[a] Çağ University, Yenice, Tarsus, Mersin, Turkey

Abstract

This study has been conducted to evaluate the transition marker (TM) usage in the doctoral dissertations written by the native speakers (NSs) of English and the Turkish speakers (TSs) of English. The purpose is to compare the TM usage in the introduction, results and discussion, and conclusion sections by both groups' randomly selected PhD dissertations in the field of ELT between the years 2010 and 2014. The WordSmith Tools 5.0 software is used in order to analyze the data. TMs were analyzed in terms of percentages, frequencies per 1,000 words and they were interpreted by calculating the log-likelihood (LL) value whether there was a significant difference in their usage. The results indicated that the frequencies, and frequencies per 1,000 words of the TM usage in the sections which were investigated of the doctoral dissertations of each group were different.

Keywords: metadiscoursive interaction; corpus; corpus-based; transition markers (tms); doctoral dissertations; Turkish speakers (TSS) of English; native speakers (NSS) of English

1. Introduction

The term "corpus" is the Latin for "body". Thus, it can be said that a corpus is any body of a text. In the language sciences, a corpus is a body of written or transcribed speech which can serve a basis for linguistics analysis and description (Kennedy, 1998). Sinclair (1994) defined "corpus" as "a collection of pieces of language that are selected and ordered according to explicit linguistic criteria in order to be used as a sample of language" (p. 2). In an earlier publication he had explained "corpus" as "a collection of naturally-occuring language text, chosen to characterize a state or variety of language" (Sinclair, 1991, p. 171). Charteris (2004) emphasized a corpus as "…any large collection of texts that arise from a natural language use; in a linguistic context, it is in contrast to other types of text that were invented specifically for illustrating a point about language" (p. 30). Hunston (2002) defined the term 'language corpus' as written, or spoken linguistic data collections, which are organized, or compiled with an aim to describe a specific pattern of a language, or present some varieties of a language.

---

* Corresponding author. Tel.: +0-324-651-4800

 E-mail address: zehrakoroglu@gmail.com

Corpus research in general has been growing since the 1960's and providing new insights into many areas of language structure and use, offering opportunities to examine the actual language use in a large scope of naturally occurring texts and to expand the scope of earlier investigations (Granger, 1998). Virtanen (1998) had pointed out that corpus analyses were useful for pilot studies. She also noted that corpus analysis methods could be used in combination with other methods of analyses for a more complete understanding and a more complex interpretation of the data. Kilimci (2001), for example, investigated the lexical profile of EFL learners through corpus query techniques. He also investigated the constructional and functional properties of prepositions in the essays by advanced Turkish EFL learners and compared them with prepositions in the writings of the native speakers of American English (NS) (Kilimci, 2002). In another study, Kilimci (2003) examined the stance and attitude in advanced Turkish learners' written discourse.

1.1. Use of Transition Markers in the Field of ELT and Written Production

In writing, developing students' ability to use transition markers (TMs) is helpful in connecting the sentences effectively as well as paragraphs, showing the logical or semantic relations between the previous information, and facilitating readers' interpretation of the whole discourse effectively (Ali et al., 2012). The use of TMs also enables speakers or writers make the context more accessible to listeners or readers and constrain their interpretation of message through using TMs in communication (Swan, 2005). Accordingly, awareness of the use and practicality of TMs can immensely contribute to the overall quality of the discourse created by English language learners. Rahimi (2011) points out that TMs constitute an essential component of communicative competence that they help learners produce fluent and meaningful discourse in English.

The present study attempts to investigate the use of TMs in the introduction, results and discussion, and conclusion sections of the PhD dissertations. The aim is to examine the similarities and the differences between the native speakers (NSs) of English and the Turkish speakers (TSs) of English in the field of ELT. The reason for the selection of the TMs as the linguistic elements to investigate in this study is their importance for the coherence and the cohesion of the academic texts. Therefore, this particular corpus-based study focuses on the question of whether the transition marker (TM) use of the introduction, results and discussion, and conclusion sections in the PhD dissertations written by the native speakers (NSs) of English and the Turkish speakers (TSs) of English indicate any difference in the field of ELT.

2. Limitations

The present study is limited to identifying the use of transition markers (TMs) in the doctoral dissertations written by the native speakers (NSs) of English and the Turkish speakers (TSs) of English in the field of ELT between 2010 and 2014. In addition, merely the introduction, the results and discussion, and the conclusion sections of the PhD dissertations were analyzed.

3. Method

The purpose of this particular corpus-based study is to investigate the TM use in the introduction, results and discussion, and conclusion sections of doctoral dissertations written by the TSs and the NSs. The TMs were analyzed in terms of frequency and log-likelihood (LL) by means of comparing the data groups. This study was designed as twofold: a descriptive study as descriptive statistics gives numerical and graphic procedures to summarize a collection of data in a clear and understandable way (Jaggi,

2003, p. 1) and the quantitative research because it is the numerical representation and manipulation of observations for the purpose of describing and explaining the phenomena that those observations reflect (Wikipedia Encyclopedia, 2015). "The obvious benefits of quantitative data are that the numerical form makes comparison easy, data are standardized, visible and amenable to the tests of classical survey statistics" (Cooper & Branthwaite, 1977 cited by Hart, 1987, p. 29).

3.1. Instrument(s)

Two different instruments were used to collect the data for this study: a) the doctoral dissertations written by the native speakers (NSs) of English and b) the Turkish speakers (TSs) of English in the field of ELT. Fifty PhD dissertations and of the NSs fifty ones of the TSs (100 theses) between the years 2010 and 2014 were randomly selected to analyze and to compare the usage of the TMs in the introduction, results and discussion, and conclusion sections.

3.2. Data collection and analysis procedures

Randomly selected 100 (50 TSs and 50 NSs) doctoral dissertations were analyzed by the researcher, by picking equally ten theses per each year. The data analyses included computer-supported tools of these two corpora. First of all, the introduction, results and discussion, and conclusion sections were extracted and saved as text files and all the other chapters were excluded from the data. Accordingly, each set of corpus was uploaded to the programme of WordSmith Tools 5.0. In this study, the usage of TMs were analyzed according to their frequency per 1,000 words. In addition to the frequency analysis, log-likelihood (LL) calculation was also used as the statistical analysis method to indicate the overuse which is referred as the higher frequency of occurrence, and the underuse which is defined as the lower frequency of occurrence for the analyzed data. When the expected relative frequency is lower than 5, most tests to measure statistical significance, such as chi-square, are unreliable, except for LL tests (Rayson & Garside, 2000, cited in Buysse, 2011).

4. Results

The data were obtained and analyzed from the introduction, the results and discussion, and the conclusion sections of the doctoral dissertations written by the NSs and the TSs considering the differences in the use of the TMs in the field of ELT. Overall Frequency Analysis of the Transition Markers in the Doctoral Dissertations Written by NSs and TSs is presented in the table below.

Table 1. Overall Frequency and LL Analysis of the TMs in the Doctoral Dissertations Written by the NSs of English and the TSs of English

	TSs	NSs	LL Value
Corpus Size	2,756,430	2,880,750	
TMs (n)	37206	34756	**+226.51***
n per 1,000	14	12	
Frequency (%)	0.1	0.1	

As observed from Table 1, the corpus size in the PhD dissertations written by the TSs of English (2,756,430) was less than the NSs of English (2,880,750). The total usage of the TMs in the introduction, the results and discussion, and the conclusion sections of the TSs' dissertations was 37206. On the other

hand, the NSs used 34756 TMs in the mentioned sections. The overall result also indicated that the TSs of English used the TMs more than the NSs of English in all three sections. The total of the TMs by the TSs (14) per 1,000 was observed to be higher than the NSs (12). However, both TSs of English and NSs of English used the TMs equally in every 100 words in the introduction, results and discussion, and conclusion sections of their doctoral dissertations.

In addition to the TM overall frequency analysis, to determine the difference between frequencies of both groups, and the significant values of overuse or underuse in the groups, LL calculation was applied. The LL value in the PhD dissertations of the TSs of English displayed a high amount of overuse as +226.51 which was statistically significant when compared to the NSs of English in all three sections. For the dissertations, the differences between the use of the TMs in two groups were compared by their frequency separately for each section.

Table 2. Frequency and LL Analysis of the TMs for the Introduction Section in the Doctoral Dissertations Written by the NSs of English and the TSs of English

	TSs	NSs	LL Value
Corpus Size	2,756,430	2,880,750	
TMs (n)	6764	6675	**+11.05***
n per 1,000	3	2	
Frequency (%)	0.03	0.02	

In Table 2, the frequency of TMs in the groups was shown by means of total number of TMs and proportion of TMs per 1,000. It was revealed that the TM usage in the introduction section for the TSs was slightly higher (6764) than those of the NSs (6675) in their dissertations. On the other hand, the total usage of the TMs of the TSs of English (3) per 1,000 words was observed to be higher than the NSs of English (2). In addition; it was indicated that the frequency of TMs in the TSs' PhD dissertations was 0.03 whereas it was 0.02 in the NSs' dissertations. According to the frequency results, the TSs of English used TMs more frequently than the NSs of English in every 100 words in their doctoral dissertations' introduction section.

LL calculation was applied to observe the difference betweeen frequencies of the NSs of English and the TSs of English, and the significant values of overuse or underuse of TMs in the groups. The LL frequency of the TMs indicated an overuse in the introduction section with an +11.05 LL value for the PhD dissertations of the TSs and there was a significant difference in the frequency of the TMs between the two groups' doctoral dissertations (PhD) in the mentioned section.

Table 3. Frequency and LL Analysis of the TMs for the Results and Discussion Section in the Doctoral Dissertations Written by the NSs of English and the TSs of English

	TSs	NSs	LL Value
Corpus Size	2,756,430	2,880,750	
TMs (n)	21054	16474	+780.65*
n per 1,000	8	6	
Frequency (%)	0.08	0.06	

In Table 3, the frequency of TMs in the doctoral dissertations (PhD) revealed that the TSs of English used 21054 TMs while the NSs of English used 16474 TMs in the results and discussion section. The

results indicated that the TMs used in this section by the TSs were higher than the NSs of the doctoral dissertations (PhD). By means of frequency per 1,000 words, while 8 TMs were used by the TSs of English, the NSs of English used 6 TMs in this section. The frequency of TMs in the results and discussion section of the TSs' dissertations revealed 0.08 TMs and the NSs displayed 0.06 TMs per 100 words. According to the frequency results, the TSs of English used higher amount of TMs than the NSs of English in every 100 words in the results and discussion section of their doctoral dissertations.

LL calculation was applied in addition to the frequency analysis. The LL value of the TMs in the results and discussion section of the PhD dissertations between the TSs of English and the NSs of English revealed an overuse as +780.65. Moreover, the results indicated that there was a significant difference in the frequency of the TMs between the two groups' doctoral dissertations in the results and discussion section.

Table 4. Frequency and LL Analysis of the TMs for the Conclusion Section in the Doctoral Dissertations Written by the NSs of English and the TSs of English

	TSs	NSs	LL Frequency
Corpus Size	2,756,430	2,880,750	
TMs (n)	9388	11607	-147.29*
n per 1,000	3	4	
Frequency (%)	0.03	0.04	

As illustrated in Table 4, it was observed that the frequency of the TMs used by the TSs of English was 9388 whereas it was 11607 by the NSs of English in the conclusion section of their doctoral dissertations. There was an overuse in the NSs of English due to the frequency difference. In addition, the total usage of the TMs of the TSs (3) per 1,000 words was observed to be less than the NSs (4) and there was a frequency difference between the groups. The frequency per 100 words in each group also indicated a difference of the TMs used between the NSs of English (0.04%) and TSs of English (0.03%).

In addition to the frequency analysis, LL calculation was applied. The LL value of the TMs in the conclusion section of the PhD dissertations between the TSs of English and the NSs of English revealed a high amount of underuse as -147.29 which was statistically significant in Table 4.

Table 5. LL Frequency of the TMs in the Doctoral Dissertations' Three Sections Among the NSs of English and the TSs of English

Sections	LL Frequency	Overused / Underused
Introduction	+11.05*	Overused in TSs
Results and Discussion	+780.65*	Overused in TSs
Conclusion	-147.29*	Underused in TSs

In conclusion, as observed from Table 5, there was an overuse in the TMs of the doctoral dissertations' introduction (+11.05%), and results and discussion sections (+780.65%) written by the TSs of English. In addition, the TSs overused the TMs statistically more significant in the results and discussion section than the introduction section of their dissertations. Moreover, the TSs' results displayed the highest overuse of TMs in the PhD dissertations' results and discussion section. However, the analysis of the conclusion section indicated an underuse for the TSs of English

(-147.29%) in their dissertations.

Considering the results given in the tables, the doctoral dissertations of the TSs of English used a wide variety of TMs in their three sections; including the introduction, results and discussion, and conclusion sections. The corpus size of the NSs of English was higher than the TSs of English. According to the overall frequency results, both groups used 0.1 TMs in every 100 words in all three sections. However, the TSs used the TMs more than the NSs per 1,000 words in terms of total TM usage. Hence, the LL overall frequency indicated the significant overuse for the TSs of English in these sections.

It is possible to state that the TM usage, including their number, frequency and usage per 1,000 words were higher for the TSs of English in the introduction, and results and discussion sections of their PhD dissertations. Among all three sections, the results and discussion section included the most frequent and the most used number of TMs in the doctoral dissertations of both groups. The high proportion of TM usage in the dissertations of the TSs could be an explanation of the significant overuse of TMs by the TSs when compared to NSs. On the contrary, the TMs had been underused by the TSs with -147.29 significant difference from the NSs because of the frequency interval of the TMs used in between the groups.

5. Conclusions

Doctoral dissertations are accepted as formal written texts produced by experts and they are required to have fulfilled academic masterpiece. Upon consideration the results of the present study, it could be revealed that the Turkish speakers (TSs) of English used a wide variety of transition markers (TMs) in their doctoral dissertations' (PhD) three sections; including the introduction, results and discussion, and conclusion sections. When the corpus size was taken account as a whole, it was observed that the size of the native speakers (NSs) of English was higher than the TSs of English. Upon the consideration of the overall frequency results, it was revealed that both groups used 0.1 TMs in every 100 words in all three sections. However, in terms of total TM usage per 1,000 words, the TSs used the TMs more than the NSs. Hence, the log-likelihood (LL) overall frequency indicated the significant overuse for the TSs of English in these sections. When the reasons for the overall results in this study were searched for, the frequency of the TMs in the PhD dissertations was observed as very important and also considered as the major factor affecting the use of TMs. Similarly, in a study conducted by Bunton (1999), the ways PhD students used metadiscourse markers in their theses were investigated and the results of his study indicated that metadiscourse markers were facilitative supporting the results of this study and raising awareness in writing academic texts.

Concerning the TM usage, including their number, frequency and usage per 1,000 words were high in the introduction, and results and discussion sections of their PhD dissertations written by the TSs of English. Among all three sections, both groups' doctoral dissertations' results and discussion section included the most frequent and the most used number of TMs. The TSs had the highest significant overuse in this section because of the text quality and the high proportion of TM usage in the PhD dissertations. In accordance with the results, the TSs' awareness might also be raised about the significance of the wide range of TMs used to establish and maintain a relationship with the readers, to express their ideas and point of view because the way they express their attitude in academic texts could be achieved through written discourse. Halliday and Hasan (1976) believed that TMs reflect the writer's positioning of one point in relation to another in creating a text. Generally speaking, TMs are the most common way of coordination and the most frequently used in academic writing (Leech & Svartvik, 1994, p. 264; Greenbaum & Quirk, 1993, p. 263). Nevertheless, the TSs of English used the TMs less than the NSs of English in the conclusion section. This significant underuse of the TMs in the PhD

dissertations' conclusion section written by the TSs could be explained because of the frequency interval of the TMs used in between the groups. As a result, the TSs of English were more attentive with their academic text production whereas the NSs of English were more convinced in their writing. Similar results were observed as Intaraprawat and Steffensen (1995) analysed the use of metadiscourse markers in persuasive essays written by English as a second language (ESL) learners. The results of their study revealed a positive relationship between the use of metadiscourse markers and learners' writing quality. In relation with the results of the study, it could be claimed that the awareness by the TSs facilitated comprehension with the presence of the TMs. It could also be suggested that the TMs have a positive effect on L2 learners' language production and seem to play an important role in the doctoral dissertations of the TSs as valid and reliable samples of academic writing.

6. Implications and Suggestions

Language learners could make use of these corpus-based materials and identify the differences of texts produced by either NSs or TSs, making up of learners' own writing could be particularly advantageous in this respect as it may offer them a higher level of engagement. Using this type of corpora in the classroom may help learners become aware of grammatical structures including in using the TMs in their writing, and also contribute to a better understanding of the issues they face in their writing advancement. In addition to the pedagogical implications, the instructors should teach conjunctive devices in complete texts rather than as isolated statements. EFL teachers could usefully present in class some academic texts with appropriately inserted TMs. They should include a variety of TMs in their writing classes within weekly schedule and they also should emphasize the TMs when teaching and in exams so that learners use the TMs accurately. Language teachers could also encourage learners to notice how TMs are used in organization of reading texts in foreign language teaching in educational settings. Moreover, a focus on the metadiscourse studies previously conducted could be included in academic writing instruction to emphasize the practices typical of native and non-native writing.

References

Ali, S., Kalajahi, R., & Abdullah, N. (2012). Discourse connectors: An overview of the history, definition and classification of the term. *World Applied Sciences Journal, 19*(11), 1659-1673.

Bunton, D. (1999). The use of higher level metatext in Ph.D. theses. *English for Specific Purposes, 18*, 41-56.

Buysse, L. (2011). The business of pragmatics. The case of discourse markers in the speech of students of business English and English linguistics. *ITL - International Journal of Applied Linguistics, 161,* 10-30.

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

Cooper, P., & Branthwaite, A. (1977). Qualitative Technology: New Perspectives on Measurement and Meaning through Qualitative Research. 20th Market Research Society Annual Conference, March, 79-92.

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

Greenbaum, S., & Quirk, R. (1993). *A student's grammar of the English language.* Essex: Longman.

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

Hart, S. (1987). The use of the survey in industrial market research. *Journal of Marketing Management*, *3*(1), 25-38.

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

Intarapraw, P., & Steffensen, M. S. (1995). The use of metadiscourse in good and poor ESL essays. *Journal of Second Language Writing, 4*(3), 253-272.

Jaggi, S. (2003). Descriptive statistics and exploratory data analysis. *Indian Agricultural Statistics Research Institute*. Retrieved May 24, 2015, from http://iasri.res.in/ebook/EB_SMAR/e-book_pdf%20files/Manual%20II/1-DescriptiveStatistics.pdf

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

Kilimci, A. (2001). Automatic extraction of the lexical profile of EFL learners through corpus query techniques. *Çukurova Üniversitesi Eğitim Fakültesi Dergisi, 21*(2), 37–47.

Kilimci, A. (2002). Constructional and functional tendencies of prepositions in the written discourse of advanced Turkish learners of English. In 11th International Conference on Turkish Linguistics-ICTL. Cyprus: The Eastern Mediterranean University.

Kilimci, A. (2003). Stance and attitude in advanced Turkish learners' written discourse. In The 38th Linguistics Colloquium: Language and Language-Processing. Piliscsaba, Hungary.

Leech, G., & Svartvik, J. (1994). *A communicative grammar of English* (2nd ed.). London: Longman.

Rahimi, M. (2011). Discourse markers in argumentative and expository writing of Iranian EFL learners. *World Journal of English Language, 1*(2), 68-78.

Rayson, P., & Garside, R. (2000). Comparing corpora using frequency profiling. In proceedings of the workshop on Comparing Corpora, held in conjunction with the 38th annual meeting of the Association for Computational Linguistics (ACL 2000). 1-8 October 2000 (pp. 1-6), Hong Kong.

Sinclair, J. (1991). *Corpus, concordance, collocation.* Oxford: Oxford University Press.

Sinclair, J. (1994). Corpus typology: A framework for classification. *EAGLES Document,* 1-18.

Swan, M. (2005). *Practical English usage.* Oxford: Oxford University Press.

Virtanen, T. (1998). Developing a linguistic corpus for philanthropic fundraising texts. *Paper presented at The Indiana Center for Intercultural Communication and Center on Philanthropy* Roundtable. Indianapolis, Indiana University-Purdue University.

Wikipedia. (2015). Text Corpus. Retrieved October 23, 2015 from http://en.wikipedia.org/wiki/Text_corpus

Ana dili İngilizce olan ve olmayanların doktora tezlerindeki üstsöylemsel etkileşimi üzerine bir çalışma

Öz

Bu araştırma anadili İngilizce olan ve olmayanların doktora tezlerindeki geçiş belirleyicilerinin kullanımını değerlendirmek amacıyla yapılmıştır. İngiliz Dili Eğitimi alanından 2010-2014 yılları arasında her bir gruptan rastgele seçilen doktora tezlerinin giriş, bulgu ve tartışma, ve sonuç bölümlerindeki geçiş belirleyicilerinin karşılaştırılması amaçlanmaktadır. Veriler WordSmith 5.0 Metin Analiz Programı kullanılarak analiz edilmiştir. Geçiş belirleyicilerinin yüzdelikleri, 1,000 kelimedeki frekansları ve kullanımları bakımından anlamlı bir farkın olup olmadığı Log-likelihood (LL) değerleri hesaplanarak yorumlanmıştır. Çalışma bulguları, her bir grubun doktora tezlerinin incelenen bölümlerindeki geçiş belirleyicilerinin kullanım yüzdeliklerinin ve 1,000 kelimedeki frekanslarının farklı olduğunu göstermiştir.

Anahtar sözcükler: Üstsöylemsel etkileşim; derlem; derlem temelli; geçiş belirleyicileri; doktora tezleri; ana dili Türkçe olanlar; ana dili İngilizce olanlar

AUTHOR BIODATA

Zehra KÖROĞLU completed her Master's and Doctoral Degrees at English Language Teaching programs of Çukurova University. Currently, she is employed as an assistant professor at Çağ University, Faculty of Arts and Sciences, the Department of English Language Teaching, Yenice, Tarsus, Mersin, Turkey.