A study on metadiscoursive interaction in the MA theses of the native speakers of English and the Turkish speakers of English

Zehra Köroğlu [a] * , Gülden Tüm [b]

[a] Çağ University, Yenice, Tarsus, Mersin, Turkey
[b] Çukurova University, Adana, Turkey

Abstract

This study has been conducted to evaluate the TM usage in the MA theses 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 MA theses 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 MA theses of each group were different.

1. Introduction

The term corpus, derived from the Latin word for body, was first encountered in the 6th century to refer to a collection of legal texts called Corpus Juris Civilis (Francis, 1992, p. 17). It has preserved this initial meaning a body of text, nevertheless, this definition is not entirely satisfactory for corpus linguists. According to one of the five definitions provided by the Oxford English Dictionary, a corpus is "the body of written or spoken material upon which a linguistic analysis is based". It results that a corpus is not just a collection of texts; it represents in fact "a collection of texts assumed to be representative of a given language, dialect, or another subset of a language, to be used for linguistic analysis" (Francis, 1982, p. 7).

A corpus-based approach could bring benefits to learners both in terms of their language awareness and possibly also in their language acquisition (Boulton, 2009, p. 37) and to see empirical descriptions

---

* Corresponding author. Tel.: +0-324-651-4800
  E-mail address: zehrakoroglu@gmail.com

of language use, identify the frequent patterns, and understand the usage of particular forms and words in different registers (Biber & Reppen, 2002). Researchers have stated that transition markers (TMs) is one of the elements used by writers to encode the message in the form of a written text and the reader must use them to interpret the message given by the writer. Using TMs appropriately and correctly in written language is an important component of textual competence.

When used adequately, they act as guideposts for the reader to perceive the text from the writer's perspective and to see shifts and changes in thought, comparisons, contrasts and countless other relational concepts (Dublin & Olshtain, 1980; Holloway, 1981; Sloan, 1983). Meyer et. al. (1980) suggested that the use of TMs helped organize discourse representation and faster discourse segment processing. As stated by Virtanen (2004), appropriate and correct use of TMs is important because they indicate the kinds of thought processes involved in the text, and they invite the reader to construe particular logico-semantic relations between units of the text. Similarly, Zamel (1983) has stated that TMs are very important in writing because they signal the relationship between ideas and affected the meaning that the writer is trying to convey. This helps understanding the thought of the writer clearly and affected the writing quality. However, a number of studies have shown that the use of TMs is problematic for foreign language learners. One reason is that TMs are not always used and that they have to be used with discrimination. The other problem is that the use of TMs is sensitive to discourse type which might cause difficulty for learners. And the last issue is that TMs usage may vary across languages and not all languages mark TMs explicitly as in English (Altenberg & Tapper, 1998). According to Celce-Murcia and Larsen-Freeman (1999), if no TM is used or if the wrong TM is used, the speaker or writer intention can be misunderstood. They have stated that the use of TMs is particularly more important in reading than speaking because there are no paralinguistic cues (e.g., tone of voice) or extralinguistic cues such as gestures to help the reader understand the writer's intention. The reason is that readers may not have any background knowledge about the subject written, or they may fail to activate their schemata even if they have some background knowledge about the subject.

1.1. Effect of Transition Markers on Sentence Processing and Comprehension

Some TMs, mostly used in formal writing styles, serve as directional guides for text receivers about how the incoming information should be interpreted and integrated with the preceding discourse segments (Halliday & Hasan, 1976; Carpenter & Just, 1977). In addition, they also provide insight about the speaker's attitude to the content of the discourse (Swan, 1980), and indicate the relative importance of the ideas in the text (Jung, 2003). Their specific meaning is determined by the context, their core meaning is procedural rather than conceptual (Fraser, 1999). In other words, while coherence relation is an intrinsic part of cognitive representations, linguistic markers are a surface code that can facilitate the process of formation of coherent text representations (Sanders & Noordman, 2000).

The present study attempts to investigate the use of TMs in the introduction, results and discussion, and conclusion sections of the MA theses. 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 of 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 use of transition marker (TM) in the introduction, results and discussion, and conclusion sections of the MA theses written by the native speakers (NSs) of English and the Turkish speakers (TSs) of English indicate any difference in the field of ELT. As highlighted by Hyland (2005), writers convey their ideas more effectively as they

allow them to take an appropriate stance and accomplish their communicative purposes with the help of the discourse markers.

2. Limitations

The present study was limited to identifying the use of transition markers (TMs) in the MA theses 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 MA theses were analyzed.

3. Methodology

The purpose of this particular corpus-based study is to investigate the use of TM in the introduction, results and discussion, and conclusion sections of MA theses 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 (Babbie & Mouton, 2001). "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. Data Collection

The data for this study was collected from: a) the MA theses written by the native speakers (NSs) of English and b) by the Turkish speakers (TSs) of English in the field of ELT. Fifty MA theses of the Turkish speakers and fifty ones of the Native American speakers (totally 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

Permission e-letters (84) were sent to personal and/or university e-mails, and personal Facebook accounts of the TSs of MA theses, and 50 permissions were obtained from them in order to analyze their MA theses. Totally 100 (50 TSs and 50 NSs) theses were randomly chosen by the researcher, by picking equally ten theses between the years 2010 and 2014. 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 use 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. Findings and Discussion

The data were obtained and analyzed from the introduction, the results and discussion, and the conclusion sections of the MA theses 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 MA Theses Written by NSs and TSs was presented in Table 1 below.

Table 1. Overall Frequency and LL Analysis of the TMs in the MA Theses
Written by the NSs and the TSs

	TSs	NSs	LL Value
Corpus Size	1,754,429	1,177,474	
TMs (n)	26805	8983	**-31.98***
n per 1,000	15	16	
Frequency (%)	0.2	0.2	

As observed in Table 1, the corpus size appears higher for the MA theses written by the TSs (1,754,429) than the NSs (1,177,474). It was revealed that the TM usage in all three sections for the TSs was higher than those of the NSs' theses. The frequency of TMs in the groups was indicated by means of total number of TMs and proportion of TMs per 1,000. However, the total use of the TMs by the NSs (16) of English per 1,000 was observed to be higher than those of the TSs (15) of English in all three sections. According to the overall frequency results, both groups used similar amount of TMs (0.2) in every 100 words in their MA theses' introduction, results and discussion, and conclusion sections.

In addition to the frequency analysis, to determine the difference between frequencies of the TSs of English and the NSs of English, and the significant values of overuse or underuse for the TMs in the groups, LL calculation was applied. The LL value of the TSs displayed an underuse as -31.98 which was statistically significant when compared to the NSs MA theses in all three sections.

Table 2. Frequency and LL Analysis of the TMs for the Introduction Section in the MA Theses Written by the NSs and the TSs

	TSs	NSs	LL Value
Corpus Size	1,754,429	1,177,474	
TMs (n)	5357	3461	**+3.06***
n per 1,000	3	3	
Frequency (%)	0.03	0.03	

In Table 2, both the corpora size and TM usage were observed higher for the TSs than the NSs in the introduction section of their MA theses. Furthermore, by means of frequency per 1,000 words, similar frequency results appeared between the TSs (3) and the NSs (3). Similarly, both the TSs (0.03) and the NSs (0.03) used the same amount of TMs in every 100 words in their introduction section of the MA theses when checked the frequency results. In addition to these results, LL calculation was applied within the groups in order to observe the difference betweeen frequencies of the TSs and the NSs and the significant values of overuse or underuse. The LL frequency indicated an overuse in the TMs of the TSs' MA theses' introduction section with an +3.06 LL value and there was a significant

difference in the frequency of the TMs between the MA theses of two groups in the introduction section.

Table 3. Frequency and LL Analysis of the TMs for the Results and Discussion Section in the MA Theses Written by the NSs and the TSs

	TSs	NSs	LL Value
Corpus Size	1,754,429	1,177,474	
TMs (n)	15670	10039	**+13.27***
n per 1,000	9	9	
Frequency (%)	0.09	0.09	

The mostly used TMs in the MA theses' results and discussion section written by the TSs of English were presented in Table 3. As indicated in the results, the TSs used TMs more (15670) than the NSs (10039). Even though the difference is the usage of MSs in both the MA theses by the TSs and the NSs, in number per 1,000 words (9) and with a frequency of 0.09 TM usage in this section were similar. LL calculation was applied in addition to the frequency analysis for the TMs. In Table 3, the LL value related to the results and discussion section of the MA theses was given and it is revealed that there was a significant difference in the frequency of the TMs between the two groups' MA theses in the results and discussion section because the TSs used overuse TMs as +13.27.

Table 4. Frequency and LL Analysis of the TMs for the Conclusion Section in the MA Theses Written by the NSs and the TSs

	TSs	NSs	LL Frequency
Corpus Size	1,754,429	1,177,474	
TMs (n)	5778	5483	**-334.83***
n per 1,000	3	5	
Frequency (%)	0.03	0.05	

As presented in Table 4, the total usage of the TMs in the conclusion section in the MA theses written by the TSs was 5775 whereas it is recorded as 5483 TMs in the theses written by the NSs. However, by means of frequency per 1,000 words, the TMs of the TSs (3) was observed to be less than the NSs (5). In addition, in the conclusion section of the MA theses, the TSs (0.03%) used TMs less than the NSs (0.05%). Besides the frequency analysis, LL calculation was applied and the LL value for the TMs in the conclusion section of the MA theses between the TSs and the NSs revealed a high amount of underuse as -334.83, which was statistically significant as indicated in Table 4.

Table 5. LL Frequency of the TMs in MA Theses' Three Sections Among the NSs and the TSs

Sections	LL Frequency	Overused / Underused
Introduction	+3.06*	Overused in TSs
Results and Discussion	+13.27*	Overused in TSs
Conclusion	-334.83*	Underused in TSs

Consequently, as clearly pointed in Table 5, there was a significant overuse of the TMs in the introduction (+3.06%), and results and discussion sections (+13.27%) of MA theses written by the

TSs. Moreover, the TSs overused the TMs statistically more significant in the results and discussion section than the introduction section of the MA theses. However, the analysis of the conclusion section in the MA theses indicated a significant underuse for the TSs (-334.83%).

Considering the results given in the tables, in the three sections of the MA theses by the TSs, it could be interpreted that a wide variety of TMs were applied. Besides, their corpus size and total TM usage were higher than the NSs of English. According to the overall frequency results, both groups used 0.2 TMs in every 100 words. There was a high overall TM usage per 1,000 words in all three sections in the MA theses of the NSs. Hence, the LL frequency indicated the significant underuse in the MA theses for the TSs in these sections.

5. Conclusion

Regarding the MA theses' three sections of which are the introduction, results and discussion, and conclusion sections written by the Turkish speakers (TSs) of English and the native speakers (NSs) of English, it could be interpreted that the TSs used a wide variety of transition markers (TMs). Their corpus size and total TM usage were higher than the native speakers (NSs) of English in all three sections. It can be stated that the difference between the TMs of the TSs of English and NSs of English in this present study might be stemming from the use of TMs in the MA thesis. Related with the overall findings, it could be interpreted that the TSs did not tend to make their aims visible in their theses and did not explicitly state their ideas through the use of the TMs.

The TM usage per 1,000 words and frequency of the TMs in the MA theses' introduction section, and the results and discussion section by the TSs of English and the NSs of English were equal. Among all three sections, the results and discussion section included the most frequent and the most used amount of TMs in the MA theses of both groups. Several studies have attempted to illustrate how conjunctions contribute to better understanding of written discourse. Some studies contended that there was a positive correlation between a number of cohesive devices and effective writing (Ferris, 1994; Field & Oi, 1992; Jin, 2001; Neuner, 1987).

The frequency of the TMs in the theses is very important and considered the major factor that affects the use of TMs. By means of frequency per 1,000 words, in the results and discussion section, the TSs used the TMs three times more than the other mentioned sections. However, the NSs used the TMs three times more in the results and discussion section than the introduction section, and approximately two times more than the conclusion section. The high proportion of TM usage in this section of the TSs could be an explanation of the significant overuse of TMs for the TSs of English when compared to the NSs of English. Supporting the studies claiming that cohesive devices affected the quality of text, Liu and Braine (2005) observed that there was a significant relationship between the number of conjunctions used and the quality of the argumentative writing created by the students. In the present study, the TSs used the TMs less than the NSs in the conclusion section. Moreover, this significant underuse of the TMs in the MA theses' conclusion section could be explained because of the frequency interval of the TMs used in between the TSs and the NSs. The NSs of English might be more cautious with their academic writing whereas the TSs of English could formally used the TMs in their writing since it is their target language. In other words, the NSs seemed to be more tentative in expressing themselves as in their native language they have no hesitation on the form of the language they use when compared to the TSs.

The results of the study suggested that with the significant use of TMs, the TSs create cohesion more in their MA theses since they apply more formal language rules or structures while producing written texts. Besides, the awareness of TM usage that contributes to the cohesiveness of the text

should help the writers recognize the links between the concepts and identify important information in the MA thesis. It could also be interpreted that the TMs were overused by the TSs could be due to the effective academic writing style they acquired or experienced during their learning process. In addition, it could be mentioned that the TSs could express the relative importance of their ideas in the theses more than the NSs since NSs naturally apply their mother tongue while establishing statements composing texts. Related to the findings, Sanders and Noordman (2000) indicated that conjunctions helped the reader construct representations. According to their study, it was explicit that the appropriate use of conjunctions contributed to the clarity and comprehensibility of a text. Upon consideration of the clarity and comprehensibility in the analysis of the MA theses, the same results are supported significantly especially in the introduction, and the results and discussion sections. Based on the results of this present study, it might be concluded that both groups' awareness of the importance of the TMs make their academic products more immersive.

References

Babbie, E. & Mouton, J. M. (2001). *The practice of social research.* Cape Town: Oxford University Press.

Biber, D., & Reppen, R. (2002). What does frequency have to do with grammar teaching? *Studies in Second Language Acquisition, 24,* 199-208.

Boulton, A. (2009). Testing the limits of data-driven learning: language proficiency and training. *ReCALL, 21*(1), 37–54.

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.

Carpenter, P.A., & Just, M. A. (1977). Integrative processes in comprehension. In D. LaBerge & J. Samuals (Eds.), *Basic processes in reading: perception and comprehension* (pp. 217-241). Hillsdale, NJ: Erlbaum.

Celce-Murcia, M., & Larsen-Freeman, D. (1999). *The grammar book: An ESL/EFL teacher's course* (2nd ed.). USA: Newbury House/Heinle & Heinle.

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.

Dublin, F., & Olshtain, T. (1980). The interface of reading and writing. *TESOL Quarterly 14,* 353-363.

Ferris, D. R. (1994). Lexical and syntactic features of ESL writing by students at different levels of L2 proficiency. *TESOL Quarterly*, *28*(2), 414-420.

Field, Y., & Oi, Y. L. M. (1992). A comparison of internal conjunctive cohesion in the English essay writing of Cantonese speakers and native speakers of English. *RELC Journal*, *23*(1), 15-28. Retrieved February 4, 2017 from http://dx.doi.org/10.1177/003368829202300102

Francis, W. N. (1982). Problems of assembling and computerizing large corpora. In S. Johansson (Ed.), *Computer corpora in English language research* (pp. 7-24). Bergen: Norwegian Computing Centre for the Humanities.

Fraser, B. J. (1999). What are discourse markers? *Journal of Pragmatics, 31*(7), 931-952.

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.

Holloway, D. W. (1981). Semantic grammars: How they can help teach writing. *College Composition and Communication, 32,* 205-218.

Hyland, K. (2005). *Metadiscourse: Exploring interaction in writing.* Bodmin: MPG Books.

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

Jin, W. (2001). *A quantitative study of cohesion in Chinese graduate students' writing: Variations across genres and proficiency levels*. Paper presented at the Symposium on Second Language Writing at Purdue University West Lafayette, Indiana.

Jung, E. H. (2003). The role of discourse signaling cues in second language listening comprehension. *Modern Language Journal,* 87(4), 562-577.

Liu, M., & Braine, G. (2005). Cohesive features in argumentative writing produced by Chinese undergraduates. *System*, *33*(4), 623-636. Retieved February 5, 2017 from http://dx.doi.org/10.1016/j.system.2005.02.002

Meyer, B. J. F., Brandt, D.M., & Bluth, G.J. (1980). Use of top-level structure in text: Key for reading comprehension in ninth-grade students. *Reading Research Quarterly, 16*(1), 72–103.

Neuner, J. L. (1987). Cohesive ties and chains in good and poor freshman essays. *Research in the Teaching of English*, *21*(1), 92-105.

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.

Sanders, T., & Noordman, L. (2000). The role of coherence relations and their linguistic markers in text processing. *Discourse Processes, 29*, 37-60.

Sloan, G. (1983). Transitions: Relationships among t-units. *College Composition and Communication, 34,* 447-453.

Swan, M. (1980). *Practical English usage.* Oxford: Oxford University Press.

Virtanen, T. (2004). Point of departure: Cognitive aspects of sentence-initial adverbials. In T. Virtanen (Ed.), *Approaches to cognition through text and discourse* (pp. 79-97). Berlin: Walter de Gruyter.

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

Zamel, V. (1983). The composing processes of advanced ESL students: Six case studies. *TESOL Quarterly, 17,* 165-187.

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

Öz

Bu araştırma anadili İngilizce olan ve olmayanların yüksek lisans tezlerindeki geçiş belirleyicilerinin kullanımını değerlendirmek amacıyla yapılmıştır. İngiliz Dili Eğitimi alanından 2010-2014 yılları arasında her bir gruptan rastgele seçilen yüksek lisans 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 yüksek lisans tezlerinin incelenen bölümlerindeki geçiş belirleyicilerinin kullanım yüzdeliklerinin ve 1,000 kelimedeki frekanslarının farklı olduğunu göstermiştir.

Anahtar sözcükler: üstsöylemsel etkileşim, derlem, derlem temelli, geçiş belirleyicileri, yüksek lisans 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.

Gülden TÜM is an associate professor at Çukurova University, Faculty of Education, the Department of English Language Teaching, Adana, Turkey. She has published articles in international and national journals.