LEXICAL BUNDLES IN LEARNER WRITING:
AN ANALYSIS OF ENGLISH ARGUMENTATIVE ESSAYS OF TURKISH AND KAZAKHSTANI ELT STUDENTS

MA THESIS

Madina YÖRÜK

Eskişehir 2021

LEXICAL BUNDLES IN LEARNER WRITING: AN ANALYSIS OF ENGLISH ARGUMENTATIVE ESSAYS OF TURKISH AND KAZAKHSTANI ELT STUDENTS

Madina YÖRÜK

MA THESIS

Program in English Language Teaching

Department of Foreign Language Education

Advisor: Prof. Dr. Gül DURMUŞOĞLU KÖSE

Eskişehir

Anadolu University

Graduate School of Educational Sciences

January 2021

JÜRİ VE ENSTİTÜ ONAYI

ABSTRACT

LEXICAL BUNDLES IN LEARNER WRITING:
AN ANALYSIS OF ENGLISH ARGUMENTATIVE ESSAYS OF
TURKISH AND KAZAKHSTANI ELT STUDENTS

Madina YÖRÜK

Department of Foreign Language Education, Programme in English Language Teaching
Anadolu University, Graduate School of Educational Sciences, January 2021
Advisor: Prof. Dr. Gül DURMUŞOĞLU KÖSE

Recent publications have demonstrated how the advantageous use of lexical bundles is crucial in the forming of spoken and written discourse. Numerous studies are available in the literature that mentions the structural and functional differences between the use of lexical bundles by native speakers of English and non-native EFL students. This study examined the use of lexical bundles in Turkish, Kazakhstani and English L1 speakers' writing, in the form of argumentative essays with the aim of unveiling the differences in the use of lexical bundles by these three groups in terms of frequency, structures and functions. In the light of the findings of this study, they confirmed the general tendency of lexical bundles tradition that native speakers demonstrate to have less multi-word expressions than non-native speakers. In terms of structural categorization, overall numbers reported that noun phrases were prevalent in all four corpora. Anadolu University students and native speakers most frequently employed referential expressions. Eskişehir Osmangazi University and Ablai Khan University students mainly preferred discourse organizers. Native speakers' functional and structural analysis yielded almost similar results with non-native students. Statistical analysis revealed that non-native speakers overused frequently recurring word sequences in comparison with native speakers.

Keywords: Lexical bundles, Corpus linguistics, Non-native speakers, Native speakers, Argumentative essay

ÖZET

ÖĞRENCİ YAZIMINDA SÖZCÜK ÖBEKLERİ:
TÜRK VE KAZAKİSTANLI İNGİLİZCE ÖĞRETMENLİĞİ BÖLÜMÜ
ÖĞRENCİLERİNİN TARTIŞMACI METİNLERİNİN ANALİZİ

Madina YÖRÜK

Yabancı Diller Eğitimi Anabilim Dalı, İngilizce Öğretmenliği Programı
Anadolu Üniversitesi, Eğitim Bilimleri Enstitüsü, Ocak 2021
Danışman: Prof. Dr. Gül DURMUŞOĞLU KÖSE

Son dönem araştırmaları tekrarlanan çok kelimeli ifadelerin veya sözcük öbeklerinin etkin kullanımının yazılı ve sözlü akademik söylemin oluşturulmasında önemli rol aldığını göstermişlerdir. Bununla birlikte, literatürde ana dili İngilizce olan ve İngiliz dilini yabancı dil olarak öğrenen öğrencilerinin sözcük öbeklerini kullanımı arasındaki yapısal ve işlevsel farklılıklardan bahseden birtakım çalışma bulunmaktadır. Bu çalışmada ana dili Türkçe, Kazakça, Rusça olan İngilizce Öğretmenliği Bölümü üniversite öğrencileri ve ana dili İngilizce olan öğrencilerin yazdıkları tartışmacı metinlerdeki sözcük öbeği kullanımı incelenecektir. Bu çalışmanın bulguları, ana dili İngilizce olanların anadili olmayanlara göre sözcük öbeklerinin daha az kullandıkları doğrulandı. Yapısal sınıflandırma açısından sonuçlar isim cümlelerinin dört derlemin tamamında en sık rastlanan türden olduğunu gösterdi. Anadolu Üniversitesi öğrencileri ve anadili İngilizce olan öğrencileri ağırlıklı olarak gönderme ifadelerini kullandı. Eskişehir Osmangazi Üniversitesi ve Ablai Khan Üniversitesi öğrencileri en çok söylem düzenleyicileri tercih ettiler. Anadili İngilizce olan öğrencilerin işlevsel ve yapısal analizi, ana dili İngilizce olmayan öğrencilerle neredeyse benzer sonuçlar verdi. Bununla birlikte, istatistiksel analiz, anadili İngilizce olmayan öğrencilerin, anadili İngilizce olanlara kıyasla sözcük obeklerini daha sıklıkla kullandığını ortaya koydu.

Anahtar Sözcükler: Sözcük öbekleri, Derlem dilbilim, Ana dili İngilizce olmayanlar, Ana dili İngilizce olanlar, Tartışmacı metin

To my husband and son,
Taner YÖRÜK and Kerem YÖRÜK

ACKNOWLEDGEMENTS

I would like to thank my professor, my family and friends for being with me throughout this long journey.

Thanks to my supervisor Prof.Dr.Gül DURMUŞOĞLU KÖSE, for being supportive, kind and friendly.

To my thesis committee, Asst.Prof.Dr. Gonca SUBAŞI and Asst.Prof.Dr. Yusuf ÖZTÜRK for many consultations, feedback, guidance and support.

Thanks to my parents, Serikbay USSEMBAYEV and Orynsha ISKENDIROVA, who always believed in me and supported me in any undertakings. My little sister Aliya ZHAKAY helped me in all aspects from the day I started this program. Without your help and encouragement, I could not have collected the data so well. I am grateful to my parents-in-law, Süleyman YÖRÜK and Huri YÖRÜK, for their care, prayers and back up.

To my brothers-in-law, Cem YÖRÜK and Metin ONAÇ, thank you for being with me since the very beginning in Bursa.

Thanks to my friends Müge TANBERK, Safiye ARSLAN ORAK and Berrin KAHRAMAN for their friendship and love. I will never forget the time we spent together.

Thanks to Res. Asst. Musa TÖMEN from Anadolu University for his continuous support and valuable feedback.

Thanks to my university friend Bethany BROBERG, for her help, friendship and beautiful time spent in Eskişehir.

Thanks to Mrs. Eva BROBERG for her help with data collection and feedback.

Lastly and most importantly, I would like to thank my lovely husband, Taner YÖRÜK for his endless support, motivation, journeys to Bursa Uludağ University, long talks about postgraduate studies and for his help during all these long years. I would have never made it without your love and support. During all my studies, my son Kerem YÖRÜK was my biggest inspiration.

Madina YÖRÜK

Eskişehir 2021

26/01/2020

STATEMENT OF COMPLIANCE WITH 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.

Madina YÖRÜK

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Page

LIST OF ABBREVATIONS

ALESS	: The Active Learning of English for Science Students
ANADOLU	: Anadolu University
BAWE	: British Academic Written English Corpus
EC	: English Corpus
EFL	: English as a Foreign Language
ELT	: English Language Teaching
ESOGU	: Eskişehir Osmangazi University
FLOB	: Freiburg-Lancaster-Oslo/Bergen Corpus
KWIC	: Key Word In Context
L1	: A Speaker's First Language
L2	: A second language is any language that a person uses other than a first or native language
LB	: Lexical Bundles
LOCNESS	: The Louvain Corpus of Native English Essays
MIT CORPUS	: Massachusetts Institute of Technology Corpus
NP	: Noun Phrase
PC	: Persian Corpus
PP	: PP
RAP	: The Research Articles in Psychology
REGEX	: Regular Expressions
SUSEC	: Stockholm University Student English Corpus
SWECCL	: Spoken and Written English Corpus of Chinese Learners
T2K-SWAL	: TOEFL 2000 Spoken and Written Academic Language Corpus
TC	: Tsinghua University Corpus
VP	: Verb Phrase
WECCL	: Written English Corpus of Chinese Learners

1. INTRODUCTION

In the last few years there has been a growing interest in the dominance of the English language as it has become the most widely used language in the academe and a worldwide method for correspondence for the spread of information and science (Öztürk & Köse, 2016). Furthermore, a recent study showed that in accordance with the educational program structure of English for Academic Purposes, figuring out how to compose scholarly texts in English is in this way turning into an undeniably significant issue for research on writing in the second language (Ruan, 2017). Phraseology assumes a vital role in this area of investigation, due to the fact Gezegin-Bal defines the concept as "the study of the structure, meaning and use of word combinations". Under this field of study are lexical bundles, idioms, and collocations. Being one of the primary phraseological varieties and a corpus accessible element of phraseology, experts define lexical bundles as the most frequent sequence of at least three words in a given register paying lesser attention to the structural status and its idiomaticity (Biber, Johansson, Leech, Conrad & Finegan, 1999). Biber, Conrad & Cortes (2004), Hyland (2008) and Wray (2008) have demonstrated how the advantageous use of lexical bundles is crucial in the forming of spoken and written discourse. Numerous studies are available in the literature that discusses the structural and functional differences between the use of recurrent word combinations by native speakers of English and non-native EFL students.

Biber and Barbieri (2007) stated that "lexical bundles are recurrent sequences of words – that are important building blocks of discourse in spoken and written registers". All recent research articles supported this argument. Additionally, Biber and Barbieri (2007) claimed that "some studies describe multi-word sequences that are idiomatic (e.g., expressions like in a nutshell), while other studies focus on sequences that are non-idiomatic but perceptually salient (e.g., you're never going to believe this)". In the present study, we concentrated on non-idiomatic expressions.

1.1. Background of the Study

In the last decade, lexical bundles have attracted much attention from research teams. Lexical bundles are repeated expressions, despite their structural status, and irrespective of their idiomaticity (Biber et al., 1999). Lexical bundles usually are not considered to be established units of language, however they might include a variety of structures, e.g. *should be noted that, in this study*. Nonetheless, lexical bundles may be

viewed as empirically derived units as they are recognized purely in accordance through their recurrence within texts. Besides, these multi-word expressions generally have perceptible roles in particular register (Allen, 2010).

Biber et al. (1999) have given sufficient data of recurrent word combinations in academic writing, such as *at the same time*, *on the basis of* and *the end of th*e. Biber and Conrad (1999) explicated that lexical bundles vary from collocations and idioms in that they are "typically no longer whole structural units, and commonly no longer constant expressions. In addition to being incomplete units (i.e. *the presence of a*), lexical bundles can combine several different parts of speech in a single string such as prepositions, nouns, or verbs. Biber and Barbieri (2007) stated that in order to be considered a lexical bundle, a recurrent word combination has to meet certain criteria about the frequency (i.e., 40 times per million words for a minimum frequency). Chen and Baker (2010) in their study selected a frequency to 25 times per million words. A different distinguishing function of lexical bundles from idiomatic expressions is that bundles (e.g. *it is difficult to, that is to say, this study is to*) are transparent in meaning from the individual words, showing the aspect of the degree of formulaicity instead of idiomaticity in language use. An additional feature is that despite being formally regular, lexical bundles (e.g. *pay more attention to, on the basis of*) are often incomplete structures (Biber et al., 1999).

Lexical bundles could be regarded as an indispensable key to reaching native-like proficiency in academic discourse. After conducting a study on lexical bundles in university spoken and written registers, Douglas Biber and Federica Barbieri (2007) stated that "these word sequences turn out to be consistently functional, indicating that high frequency is a reflection of pre-fabricated or formulaic status" and Qin (2013) believed that lexical bundles play a vital role in communication. Previous studies have almost exclusively focused on non-native novice language users on their way toward reaching a native-like written performance. Many authors as Rica-Peromingo (2009), Nekrasova (2009), Chen and Baker (2010), Wei and Lei (2011), Tenuta, Oliveira and Orfano (2012) and Bychkovska and Lee (2017) in the implication part of their studies suggested further research with the English language learners of different backgrounds. Research works of Öztürk (2014), Uysal (2012) and Güngör and Uysal (2016) have undoubtedly made a contribution to better understanding of the natural second language usage of Turkish students. Muşlu (2018) even compared Japanese and Turkish EFL learners. Moreover, Karabacak and Qin (2013) shared the data on Turkish, Chinese and

American university students. It is important and unique, because unlike Muşlu's work, this research will be based on comparing the papers of multilingual L2 learners of multiannual universities of two countries with Turkic backgrounds, who had the same type of assignment and time for its completion along with the essays of native speakers.

1.2. Statement of the Problem

Publications on the use of lexical bundles by native and non-native speakers in last decades demonstrated the importance of lexical bundles in corpus linguistics (Chen & Baker, 2010). These publications focused on different genres, L1 and L2 speakers, registers and levels of language. Hyland (2012) believed that the appropriate usage of such n-grams demonstrates the level of language proficiency and is a vital component of native-like fluency. Previous research reported that non-native speakers use less recurrent word combinations (Howarth, 1998, Erman 2009) and less varied ones, compared to native speakers (Lewis, 2009; Granger, 1998).

With respect to the findings above, in the 21st century, a period of electronic publishing, it is important for writers to produce decently-written papers. Further research is required to indicate their proper usage of lexical bundles in comparison with native colleagues. Cortes (2002), who published a great deal of research on lexical bundles, claimed that it is essential to examine the use of lexical bundles in student writing, in terms of discipline and language proficiency. Muşlu (2018) conducted research on Japanese and Turkish speakers, Biber and Ventura investigated Korean and Spanish speakers; however, there are no studies on bilingual speakers from Kazakhstan whose second mother language is Russian. Although Shin's (2018) study reported that frequency, structure and function of lexical bundles of non-natives and native speakers do have similarities, Dontcheva-Navratilova (2012), Ruan (2017) and Bychkovska and Lee (2017) reported varying findings.

1.3. Aim and Research Questions

With respect to the aforementioned studies and issues, this research aims to examine the use of lexical bundles by native English speakers, and Turkish and Kazakhstani EFL learners in terms of frequency, structures and functions.

1. What are the lexical bundles that are frequently used, and shared, by native English speakers and Turkish and Kazakhstani EFL learners?

2. What are the similarities and differences in the use of lexical bundles by these three groups in terms of frequency, structures and functions?

1.4. Significance

Hyland (2012) strongly believed that multi-word expressions are essential for instructors and learners of academic English in terms of pedagogy and language acquisition. Pang (2010) claimed that recurrent word combinations facilitate non-native writers' ability to express themselves and make sure that they are understood in the right way. Researchers focused on variety of backgrounds (Muşlu, 2018; Biber & Ventura, 2010; Juknevičienė, 2009; Wei and Li, 2011). This research presented one more comparison of lexical bundle usage from different L1 background. Consequently, this study shed light on the use of English lexical bundles not only by Turkish speakers, but also speakers of the Kazakh and Russian languages.

1.5. Limitations

The limitation of this study was a relatively small number of essays. Having higher amount of essays would help with the generalization of the results. Apart from that, collecting argumentative essays as a part of a course, not on a voluntary basis would have motivated students to write more accurate essays. Lastly, all non-native essays were written on the same topic while native essays had various topics. Ideally it would be collecting the same essays from native group of students and examining them accordingly.

1. LITERATURE REVIEW

2.1. Definition of Lexical Bundles

Different scholars referred to lexical bundles in various ways. Such as clusters (Hyland, 2008a; Schmitt, Grandage & Adolphs, 2004), recurrent word combinations (Altenberg, 1998; De Cock, 1998), phrasicon (De Cock, Granger, Leech, & McEnery, 1998), n-grams (Stubbs, 2007a, 2007b), lexical bundles (e.g., Biber & Barbieri, 2007; Cortes, 2002) are explored. Different terms "have been used in applied linguistics to denote the concept of formulaic language, such as sentence stems, prefabs or lexical phrases, formulaic sequences and lexical bundles" (Karabacak & Qin, 2013). The initial reference to lexical bundles was mentioned meticulously twenty years ago in the Longman Grammar of Spoken and Written English (Biber et al., 1999). And the description of the term denoted "bundles of words that show statistical tendency to co-occur" and "recurrent expressions, regardless of their idiomaticity, and regardless of their structural status" (Biber & Conrad, 1999). The other scholars, such as Biber, Reppen and Byrd (2002), Biber, Conrad, Cortes (2004), Biber and Barbiery (2007), Hyland (2008), Rica-Peromingo (2009), Allen (2009), Bal (2010), Qin (2013), Öztürk (2014), Grabowski and Jukneviciene (2018), Bychkovska and Lee (2017) and many others were reliant on the work of Biber et al. (1999). In general, lexical bundles were described as persistent recurrent structures that are hardly ever constituted structural units, moreover they were wide-ranging collocations: expressions of three or more units of language that co-occur. In the first and subsequent research works, it was evidently illustrated that the frequency of words in lexical bundles ranges from three to five. Nobody described lexical bundles as two-word expressions, except for two academicians Adel and Erman (2012). Their study mostly focused on four-word lexical bundles, but as they maintained "three- and two-word bundles" had also been involved in their investigation. They found that such phrase fragments as *sort of, kind of* are two-word lexical bundles. There were also examples of the most pervasive grammatical category of lexical bundles in David Allen's (2009) work - the noun phrase + of structure, *the temperature of the* and *the length of the*. These examples were identified as four-word lexical bundles, but these phrases could easily become two-word bundles (when viewed from the perspective of Adel and Erman) without the article "the". In The Longman Grammar of Spoken and Written English (1999), upon which a number of authors rested their implications, it was stated explicitly

that a lexical bundle has to include three or more words. In addition, it was indicated that two-word combinations are to be considered as a kind of lexical bundles only if they are contracted. To cite one example, the contraction “I don’t” has three lexical units “I do not” and still they are accepted as single words.

Biber, Conrad and Cortes (2004) classified lexical bundles in terms of structure and function. The structural aspect of lexical bundles was classified by them into three fragments:

1) Verb phrase fragment: is going to be, what do you think, do you want to, etc.

2) Dependent clause fragment: if you want to, I don’t know why, that there is a, etc.

3) Noun or prepositional clause fragment: a little bit about, the end of the, the way in which, etc.

Biber and Barbieri (2007) divided lexical bundles into three groups according to their discourse functions and described them by saying “they are stance bundles, discourse organizers and referential expressions. Stance bundles express attitudes or assessments of certainty that frame some other proposition. Discourse organizers reflect relationships between prior and coming discourse. Referential bundles make direct reference to physical or abstract entities, or to the textual context itself, either to identify an entity or to single out some particular attribute of the entity as especially important”.

All of the postliminary research works of such scholars like Biber and Barbieri (2007), Adel and Erman (2011), Staples and Egbert, (2003), Biber and McClair (2013), Qin (2013), Öztürk (2014), Uysal (2016), Bychkovska and Lee (2017) proceeded from it. The above identified three groups of lexical bundles in terms of function: stance expressions, discourse organizers and referential expressions. However, despite mentioning the same authors (Biber et al., 2004), the work of Suethanapornkul (2009) classified four functions of lexical bundles, the fourth one was called “special conversational functions”, which he then described as “no category” bundles in functional classification table. As an example of special conversational functions of lexical bundles, Suethanapornkul (2009) provided two expressions: *taking place in my* and *to go to school*. Meanwhile Ken Hyland (2008) proposed his own functional taxonomy of lexical bundles stating that they are similar to Biber’s (2006) three main categories, however he suggested that his classification was more research-focused and relevant to academic writing. It was

then mentioned in the recent work of Pan, Reppen and Biber (2015) that Hyland's threefold differentiation demonstrated similarities to the framework of Biber et al. (2004):

- Referential bundles as Research-oriented
- Discourse Organizers as Text-oriented
- Stance bundles as Participant-oriented

Güngör and Uysal (2016) implemented in their study Salazar's (2014) functional classification, an upgraded category of Hyland's (2008) proposal. Brief overview through Danica Salazar's book have not revealed what exactly improved in Hyland's classification, as it has the same table of lexical bundles structures. Biber, Conrad and Cortes (2004) in their analysis stated that stance bundles, especially personal, happen to appear in one of the university instructional registers of formulaic classroom teaching. This confirms the fact that every learner hears stance bundles through conversations on English lessons and as a consequence of a constant repetition of such phrases as *you have to*, every single one of them is able to use them "naturally" at the level of intuition.

There was one unusual finding in Suethanapornkul's (2009) work, who conducted the research on lexical bundles which were collected and processed through the computer program. Only one writing of learner included stance expression example (*are more likely to*) was found and that is why it was excluded from the general analysis. Unusual, because Muşlu (2018) for instance, in her paper found many examples of stance bundles in learners' argumentative essays taken from three different corpora. One of her discoveries was that "EFL learners used personal uncertain stance bundles" more than native speakers. Her other discovery was that unlike Japanese learners, Turkish learners' use of epistemic devices is closer to the native speaker use. And the results of Bychkovska and Lee (2017) proved that in the works of second language education, learners use a considerably larger number of stance bundles than native speakers. The acquisition of Suethanapornkul's (2009) data was carried out at University of Hawaii at Manoa, based on placement tests on language skills, excluding speaking tests. Such indication of stance bundles the author describes as "the lack of any knowledge of stance expressions or the inability to put the knowledge" to use (Suethanapornkul, 2009).

One more allusion to two-word lexical bundles can be noticed in one of the former studies after Adel and Erman (2012), made by the researcher named Sakol Suethanapornkul (2009). In his work, he described the way in which in general novice writers, regardless of their being native or non-native language learners drew on lexical

bundles in “fewer grams (e.g. two or three words), instead of four-gram bundles”. The characteristics of two-word lexical bundles again have the effect on its quantity index and are supposed to go thoroughly into a question: Can lexical bundles be made of two words? Biber et al. (1999), in their book “Longman Grammar of Spoken and Written English”, stated that “a lexical bundle is defined here as a recurring sequence of three or more words”. In one of the subsequent works by Adel and Erman (2012) it is written that lexical bundles can have two-word bundles. Moreover, according to Biber (2006, 2007) and Hyland (2008) taxonomies, lexical bundles cannot contain only two words. To clarify this issue, a researcher contacted Douglas Biber and he clarified that two-word bundles can be considered as lexical bundles. There was nothing wrong with it; however, he and his colleagues regarded them as being collocations.

Hyland (2012) had numerous works on lexical bundles and he examined different disciplines. According to Hyland, every discipline has its own lexical bundles and if a writer is competent in his area, he should be able to use lexical bundles properly. Otherwise, something will be missing in his works. Hyland (2008) and Biber et al. (1999) divided lexical bundles in some grammatical categories. They are: NP +of, Other NPs, Prepositional Phrase+of, Other Prepositional Phrase, Passive+Prepositional Phrase/ That Complement, Antivipatory it + V/Adj, Be + N/Adj Phrase and Others. Allen (2009) described noun phrases as the largest grammatical category of lexical bundles, e.g. *the length of the, the purpose of the, the temperature of the*. Biber et al. (1999) also mentioned that the massive part of lexical bundles is noun phrases. For instance, the findings of Hyland (2008), who investigated lexical bundles in different disciplines, showed that noun phrases frequency is higher than others. Such as 23.7% in the biology and 22.3% in the electrical engineering field.

Passive forms of lexical bundles were examined by Hyland (2008), Bibel et al.(1999) and Allen (2009). Hyland’s study illustrated that 30% of the lexical bundles are passives + prepositional phrases while Allen’s study showed only 6%. For instance, *can be said that, it is well known* and *is known that the*. Allen (2009) examined ALESS corpus in his research. His findings showed that the most used type was epistemic stance, such as *it is known that* and *is widely known that.* This epistemic stance illustrates the writer’s opinion, writer’s approval. If the writer had another opinion, the bundles like *it has been suggested, some have argued* could be used.

Biber and Barbieri (2007) described that Biber et al. (2004) and later other scholars examined the functions of lexical bundles in terms of three following types: referential expressions (e.g. students must define and constantly refine *the nature of the* problem), discourse organizers (e.g. *What I want to do is* quickly run through the exercise) and expressions of stance (e.g. *I don't know what* the voltage is here). Lexical bundles were examined in different ways. For example, students and professional writers were mentioned in works of Cortes (2004) and Hyland (2008a), native and non-native speakers were mentioned in works of Chen and Baker (2010), Cock (2000), Romer (2009), different registers like classroom discourse and textbooks by Biber et al. (2004) and Biber and Barbieri (2007).

2.2. Corpus

Written and spoken materials can be accessible through electronic database and the size of these texts can be 50,000 words or may include million words (McCarthy, 2004). The materials in corpora are taken from newspapers, magazines, books. Before downloading the materials, they all need to be scanned. Corpora can include all publications of an author. For example, Tolstoy, Hemingway, etc. The corpus gives us a chance to view every stage of a language, how it was used centuries ago with examples.

Another section of the corpus is spoken part, which includes spoken language records. They are: radio broadcasts, business meetings, recorded conversations on different subjects, TV shows, phone calls, etc. Together with the written corpus, spoken corpus helps us to see the language used in different situations and in real life. Every corpus has its own aim. For instance, one corpus is designed to make a dictionary, since for the dictionary millions of words are needed. A corpus with business meetings could help instructors to teach business class. For a professor at the university, lectures and seminars from the corpus could help to write an academic book.

According to McCarthy (2004), the use of corpus is quite easy. For instance, to find a word, it is enough to write it in the "search" box like in Google, Yandex, etc. A person might find answers to the following questions:

"What are the most frequent words and phrases in English?

What are the differences between spoken and written English?

Which tenses do people use most frequently?

What prepositions follow particular verbs?

How do people use words like can, may and might?

Which words are used in more formal situations, and which are used in more informal ones?

How often do people use idiomatic expressions and why?

How many words must a learner know in order to participate in everyday conversation?

How many different words do native speakers generally use in conversation?" (McCarthy, 2004).

Before the invention of corpus, it was challenging to see or hear the recordings or written works of writers. However, now in five minutes, it is possible to see and hear the original transcript via internet (McCarthy, 2004).

McEnery and Hardie (2011) claimed that corpus linguistics is a completely different area of linguistics. It is not straight-out related to the language, but also procedures and methods to learn the language. Not all procedures are developed till the end; however, professionals are working on improving them. It is worth mentioning that the concordance tool is developed enough, and it is the most used tool in the corpus. McEnery and Hardie (2011) believed that these developments initiated a corpus-based approach in different topics of linguistics. It is worthwhile to mention The Brown Corpus, as it the one of the oldest corpora nowadays. This linguistics corpus from the 1960's is an achievement of Henry Kucera and Nelson Francis, including one million words of 500 English text samples from 15 different genres.

2.3. Corpus-based Studies

According to Güngör and Uysal (2016) misuse, underuse and overuse of recurrent word combinations are popular among corpus studies nowadays and they are associated with non-native speakers of English. Güngör and Uysal (2016) investigated English research articles of native and non-native scholars and looked at structural and functional characteristics of lexical bundles. The data showed that L1 speakers use lexical bundles differently than L2 speakers. Native speakers preferred to use noun and prepositional phrase-based lexical bundles more often than clausal bundles. Whereas non-natives overused clausal and verb phrase lexical bundles. Another important aspect is cut-off

points of lexical bundles. Güngör and Uysal (2016) wrote that cut-off points were selected according to the size and mode of a corpus and ranged between 10 and 40 occurrences per million words. Moreover, the corpus consisted of numerous lexical bundles and was not helpful for the research. Güngör and Uysal (2016) stated that as a solution, Biber and his colleagues suggested a structural taxonomy. After a while, Biber et.al (1999) and Hyland (2012) advised to categorize lexical bundles according to their discourse functions.

Kashiha and Heng (2013) investigated the corpus of 24 academic lecture transcripts taken from BASE corpus. BASE is an online academic corpus that consists of seminars and lectures of four fields of studies. The data examined in this research are lectures across two broad fields of sciences: Hard sciences (HS) and soft sciences (SS). Soft sciences included lectures in law, politics, and CELTE (Center for English Language Teacher Education). Lectures in hard sciences were chemistry, computer, and engineering. The disciplines were selected on the basis of matching word count in each science, in contemplation of having equivalent data. Results showed that hard disciplines offered a larger number of lexical bundle use compared to soft disciplines. Prepositional and noun phrase fragments were commonly-seen structures in soft science corpus, while these structures were the most prevalent in hard sciences corpus. Results reported a higher number of referential expressions in the soft science lectures, whereas lecturers in hard sciences showed a greater tendency towards the use of discourse organizers. Wei and Li (2011) explored PhD dissertations of Chinese speakers and compared them with journal articles written by professionals. After Wei and Lei (2011), Öztürk (2014) and Güngör and Uysal (2016) investigated dissertations and journal articles, an additional study was done by advance Turkish EFL learners.

Mahlberg (2007b) investigated literary and academic texts. Her research ended up with new five functional categories of literary texts. She believed that these categories are similar to the Charles Dickens model.

"1. Labels – contain the names of characters (e.g. *Mr. Pickwick and his friends*) and names of places (e.g. *a tavern of a dropsical appearance*).

2. Speech clusters – contain a first or second person pronoun or possessive, which is taken as an indication of interaction (e.g. *what do you mean by*).

3. As If clusters – contain clusters that starts with *as if* (e.g. *as if he would have*).

4. Body Part clusters – contain at least one noun referring to a part of the human body (e.g. *his hands in his pockets*).

5. Time and Place clusters – contain a nominal time and place expression with or without a preposition (e.g. *on the top of his, the opposite side of the*)".

Mahlberg preferred to use five-word lexical bundles and she believed that shorter clusters were difficult to characterize and the occurrences were higher and flexible. She chose five-word lexical bundles because their numbers were sufficient to categorize and they appear in various texts (Mahlberg, 2007b).

2.4. Studies on lexical bundles

Table 2.1 illustrates various studies on recurrent multi-word expressions, such as name of the study, year, objective, corpus information and findings.

Table 2.1. *Recent studies on lexical bundles*

No	Study	Year	Title	Objective	Corpus	Findings
1	Douglas Biber and Federica Barbieri	2007	Lexical bundles in university spoken and written registers	Investigate the use of lexical bundles in university registers.	TOEFL 2000 Spoken and Written Academic Language (T2K-SWAL) Corpus.	Study reported that lexical bundles are frequently used in non-academic discourse and in written course management.

Table 2.1. (Continue) *Recent studies on lexical bundles*

No	Study	Year	Title	Objective	Corpus	Findings
2	Douglas Biber, You Jin Kim and Nicole Tracy-Ventura	2010	A Corpus-driven approach to comparative phraseology: lexical bundles in English, Spanish, and Korean	Investigate the extent to which formulaic language is a universal of discourse and learn whether formulaic sequences are distributed in similar ways and used for similar functions in languages other than English.	English: conversation, academic prose. Spanish: sociolinguistic interviews, academic prose. Korean: academic prose and conversation.	The distribution and nature of lexical bundles vary in the three languages. While discourse in English (especially in conversation) and Spanish (especially written prose) is composed of lexical bundles, there are few comparable sequences of recurrent words in Korean and lexical bundles have a much more peripheral role in the construction of discourse in that language.
3	Yu-Hua Chen and Paul Baker	2010	Lexical bundles in L1 and L2 Academic Writing	Identify frequently-used multi-word expressions in academic registers.	Two corpora: the Freiburg-Lancaster-Oslo/Bergen (FLOB) corpus, and the British Academic Written English (BAWE) corpus.	Two groups – native and non-native speakers show almost similar results. Both groups used discourse organizers and verb phrase bundles more than native scholars.
4	Annelie Ädel and Britt Erman	2012	Recurrent word combinations in academic writing by native and non-native speakers of English: A lexical bundles approach	Investigate the use of lexical bundles in advanced learner writing by L1 speakers of Swedish and in comparable native-speaker writing.	One million word Stockholm University Student English Corpus (SUSEC).	This study reported that native students use more complex lexical bundles and negations.
5	Olga Dontcheva-Navratilova	2012	Lexical bundles in academic texts by non-native speakers	Investigating the use of lexical bundles in non-native speaker academic discourse.	Czech university English Department students' dissertation corpus.	The students from English Department used less varied lexical bundles comparing to expert writing. The level of foreign language influences the use of lexical bundles. The higher level is more varied lexical bundles.

Table 2.1. (Continue) *Recent studies on lexical bundles*

No	Study	Year	Title	Objective	Corpus	Findings
6	Zhoulin Ruan	2016	Lexical bundles in Chinese undergraduate academic writing at an English medium university	Examine lexical bundles in academic register of Chinese students.	Essays of Chinese students from Sino-UK English university.	Since the essays were from different years of university studies of the same students, the results showed that gradually students started to use lexical bundles more frequently.
7	Rajab Esfandiari and Fatima Barbary	2017	A corpus-driven study of lexical bundles between English writers and Persian writers in psychology research articles	Compare native-English and non-native-English writers in use of lexical bundles in writing research articles structurally and functionally.	Two corpora: Persian corpus (PC) and English corpus (EC).	Writers of both groups showed a frequent use of prepositional phrase and noun phrase bundles. Findings showed that Persian speakers misuse, overuse and underuse lexical bundles.
8	Tetyana Bychkovska and Joseph J. Lee	2017	At the same time: Lexical bundles in L1 and L2 university student argumentative writing	Compare the use of lexical bundles in English argumentative essays of Chinese and English students.	105 essays of Chinese students and 101 essays of English students.	Results showed completely different usage of lexical bundles. Chinese students used more verb phrase bundles while English students preferred noun and prepositional phrase bundles.
9	Yanfeng Yang	2017	Lexical bundles in argumentative and narrative writings by Chinese EFL learners	Investigate lexical bundles in argumentative and narrative writings by Chinese EFL learners.	WECCL (Written English Corpus of Chinese Learners).	Narrative essays showed higher frequency of four-word lexical bundles than argumentative essays. Argumentative writings have more stance bundles while narrative writing have more referential expressions.

Table 2.1. (Continue) *Recent studies on lexical bundles*

No	Study	Year	Title	Objective	Corpus	Findings
10	Fan Pan and Chen Liu	2019	Comparing L1-L2 differences in lexical bundles in student and	Compare L1-L2 differences in the use of lexical bundles in master dissertations and research articles	Corpus of 2.7 million words including master dissertations and academic papers.	L2-English academic writers employ more bundle types and tokens than L1-English academic writers regardless of levels of expertise. The higher English proficiency level illustrates more frequent use of lexical bundles. Results are highly different in terms of function and structure.
11	Xiaofei Lu and Jinlei Deng	2019	With the rapid development: A contrastive analysis of lexical bundles in dissertation abstracts by Chinese and L1 English doctoral students	Compare the use of lexical bundles in thesis abstracts of Chinese and English PhD students in terms of structure and function.	PhD dissertations from the Tsinghua Corpus (Tsinghua University) and the MIT corpus (Massachusetts Institute of Technology).	Chinese students have limited repertoire of lexical bundle usage comparing to L1 English students.
12	Yu Kyoung Shin	2019	Do native writers always have a head start over non-native writers? The use of lexical bundles in college students' essays	Comparing the use of lexical bundles in academic writing of L1 and L2 speakers.	6630 essays of college students	L1 and L2 speakers illustrate similar usage of lexical bundles, such as frequently used verb phrase bundles, stance bundles and prepositional phrase bundles.

Yu-Hua Chen and Paul Baker (2010) conducted a research on lexical bundles of native and non-native student academic writing and professional writing. They examined two corpora: the Freigburg-Lancaster-Oslo/Bergen (FLOB) corpus and the British Academic Written English (BAWE) corpus. They found that L1 and L2 students both preferred using verb phrase bundles and discourse organizers. Contrarily, the authors of published academic texts employed more noun phrase bundles and referential markers.

The results showed a noticeable difference between lexical bundles used by students and professionals. A bundle like *all over the world* was overused by nonnative students while *in the context of* bundle was underused by both student groups.

Another comprehensive study which worth to mention is the study of two Swedish professors Adel and Erman (2012) which is based on comparison of lexical bundles produced by Swedish and native speakers. One million corpus of student writing named the Stockholm University Student English Corpus (SUSEC) was used in this study. The comparison revealed that native speakers had a wider range of lexical bundles and one more confirmed the previous studies that reported that non-native speakers' usage of lexical bundles is less varied. These results supported the hypothesis of Biber et al. (2004) and Hyland (2008) that "lexical bundles are the important building blocks of effective academic discourse". Meanwhile Yoon and Choi (2015) investigated lexical bundle usage by Korean university students' essays and essays of native students. Four-word bundle research showed that Korean students employed conversational bundles, such as stance bundles and personal pronouns. *On the other hand,* was the most frequent bundle; however, this bundle was used incorrectly. Ruan (2016) conducted a study on lexical bundles with Chinese university students from 1st to 4th year. Results revealed that students started to use more lexical bundles when they passed to the next study year. Functional analysis ended up with discourse organizing bundles used the most, especially by final year students. Ruan believes that the higher proficiency levels the more lexical bundles are used by students.

Staples et al. (2013) investigated lexical bundles in writing part of different levels in the TOEFL test. It is well known that in this test, there is a specific amount of time for the test taker. The results yielded that the test takers with lower score preferred bundles more frequently, compared to others. The frequency number of functional distribution illustrated that the referential bundles were used the least, and the number of stance bundles and discourse organizers were almost the same on all levels.

Bychkovska and Lee (2017) like Ruan investigated lexical bundles of Chinese and native students. 101 argumentative essays of native students and 105 argumentative essays of Chinese students. Four-word bundles were examined in terms of function and structure and the results revealed that Chinese students used more lexical bundles than native in contrast to the studies above. Native students used noun and prepositional phrase bundles and Chinese students most frequently used verb phrase bundles. Moreover,

Chinese students preferred stance bundles. The analysis indicated that Chinese students have many grammatical mistakes and improper use of prepositions and articles. Bychkovska and Lee likewise Chen and Baker suggests to include lexical bundle learning in the curriculum. They believe that rather than learning from native professionals' publications, the essays of native students with decent proficiency level from BAWE corpus for instance, could enhance the learning process.

Yang (2017) analyzed the use of lexical bundles by Chinese students by their narrative and argumentative essays. The essays were retrieved from Written English Corpus of Chinese Learner and Spoken and Written English Corpus of Chinese Learners. KfNgram sorted out the lexical bundles and functional and structural analyses were performed by Biber's two taxonomies. The results indicated that argumentative essays had higher number of four-word bundles comparing to narrative essays. Structural analysis did not reveal a major difference in both essay types. Students used referential expressions mainly in the narrative essays and stance bundles in the argumentative essays. In this study an important moment is that the genre of two writings is different. One writing is about description of something and not using referential expressions is inevitable in this type of essay. Another writing is about expressing a point of view and similarly stance bundles stand first in the functional classification. Consequently, it can be concluded that these findings were not unexpected.

Another study with the results where non-natives rely on lexical bundles more than natives is Muşlu's (2018) study with argumentative essays of natives, Japanese and Turkish EFL learners. This study was a part of her PhD research and was aimed on structural and functional characteristics of lexical bundles. Argumentative essays were retrieved from Japanese International Corpus of Learner English (JPICLE), Louvain Corpus of Native English Essays (LOCNESS) and Turkish International Corpus of Learner English (TICLE). Categorization was performed using Biber's taxonomy and WordSmith Tool was used to find three or four-word stance lexical bundles. The findings indicated that non-native learners used more lexical bundles than native speakers with high frequency of overused bundles. According to the analysis the Japanese speakers were the group that most frequently used lexical bundles. All groups of speakers used verb phrase bundles more than the other bundles.

Juknevičienė and Grabowski (2018) are authors of the research with Lithuanian and Polish students. They compared the writings of Lithuanian and Polish students with the

writings of native speakers. The essays of Lithuanian and Polish students were taken from subcorpora of International Corpus of Learner English (ICLE), namely PICLE and LICLE corpora (Polish Learner English and Lithuanian Learner English). These two subcorpora consist of undergraduate Polish and Lithuanian student writings with higher proficiency of English and these students were from linguistics departments. Native speakers' essays were retrieved from LOCNESS corpus (Louvain Corpus of Native English Essays) which is a collection of works of American and British students. The results illustrated that Polish and Lithuanian students have many similar characteristics and different usage of lexical bundles. The reason for this can be their L1 languages.

The last study on lexical bundles of learner writing is Shin's (2019) research on native and Korean novice writers. First year Korean university students wrote 6630 argumentative essays for the period 2009-2012. The Korean corpus consists of 1.6 million words. All of them had the same amount of time (50 minutes) to write an essay in a class. Correspondingly students at one university in the United States were asked to write an essay in the same amount of time. The essays of the students with different L1 background were excluded from the corpus. This corpus consists of 1414 essays around 490,610 words. The results revealed that both groups in that study had almost the same results, probably because they all were freshmen. However, non-natives used noun phrase bundles as subjects instead of objects. These learners used more colloquial expressions and academic-register features compared to native speakers. For instance, native speakers employed 70 stance bundles and non-native employed 71. Native speakers used 59 referential expressions and non-natives 60. The findings do not support Adel and Erman's (2012) claim that professional writers are "more mature academic writers" as they used more varied lexical bundles.

Cortes (2004) conducted a research on lexical bundles of professional writers and students. History and biology journals and student writings from the same field were examined in terms of structure and function. Two journals of history with 92 authors had 966,187 words and three journals of biology with 199 authors had 1,026,344 words. Student writing in history had 493,109 words and biology corpus had 411,267 words. The results indicated that students almost do not use four-word lexical bundles. Moreover, the analysis showed that the bundles that come across in the student writings were not used by expert writers.

Douglas Biber, who has many publications on lexical bundles, conducted a research with Federica Barbieri (2007) on lexical bundles in spoken and written registers. These are class management talks, written syllabi, student advising, office hours, etc. The findings indicated that recurrent word combinations were frequently used in informal registers than in academic registers. There is a contradiction, as previous research revealed that recurrent word combinations are prevalent in spoken registers, however this study confirmed that they are commonly used in written registers.

Another important name after Douglas Biber is Ken Hyland. Hyland (2008) investigated 3.5-million-word corpus of PhD dissertations, MA theses and journal articles in terms of structure, function and form. Most frequent bundles were *on the other hand, at the same time, in order to* and *in terms of*. All data in the research was from four disciplines: microbiology, electrical engineering, business studies and applied linguistics. PhD dissertations and MA theses were taken from students of five universities in China and their mother language was Chinese. 120 journal articles were from the most popular 30 journals of the four disciplines. Hyland's findings showed that even lexical bundles are not the core focus in the texts; however, they help to divide the text into different disciplines. Hylands suggested teaching learners lexical bundles using different techniques like item identification and matching.

As Zamel (1998) noted in his paper, the English language is a language of communication worldwide. However, students and researchers whose native language is not English faced challenges as they have to publish their works in English. Sometimes it is not easy to transfer all knowledge from one's native language to English (Zamel, 1998). Kumaravadivelu (2001) once wrote that he struggled for long years to be accepted by scholars whose native language is English because his L1 language was different.

2.5. Antconc Program

Anthony (2004) described the Antconc Program as "a freeware, multi-platform, multi-purpose corpus analysis toolkit, designed specifically for use in the classroom". Laurence Anthony is the author of this program. The n-gram tool sorted all lexical bundles and showed how many types and tokens each essay has. This software was convenient and freely accessible, compared to other tools. This program also allows the user to enter the number of lexical bundles needed. From the frequency tool, a user can select three or four lexical bundles. In addition, the concordance feature helps the user to immediately

see the lexical bundle within the sentence. By means of this program, it is easy to set the cut-off points in the analysis part.

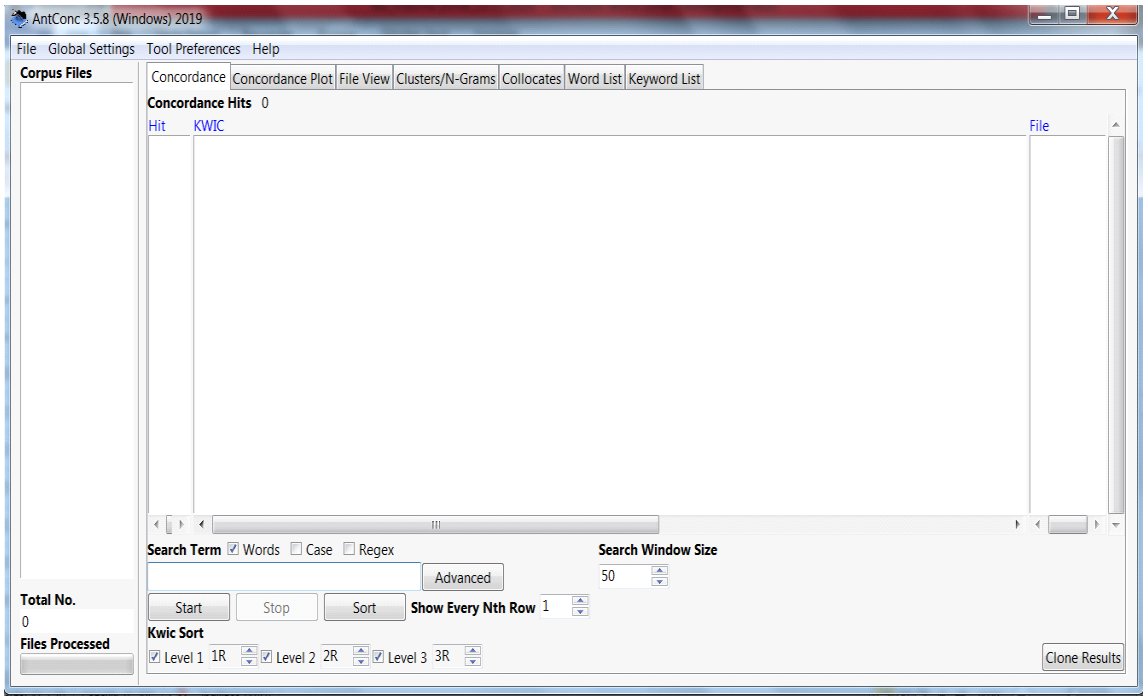

Figure 2.1. *Antconc 3.5.8 Home Screen*

Almost all analysis programs have a concordancer tool as a main tool. The concordancer tool enables acquisition of collocations, writing styles, grammar and vocabulary together with learning a foreign language (Sun & Wang, 2003). As seen in the figure below, the central screen displays all operations, and it is easy to find a command without opening tabs and the scroll down menu. Lonfils and Vanparys (2001), in their "How to design user-friendly CALL interfaces" paper, stated that additional windows and scroll down menus could complicate the usage of the program. The tabs of the program are designed as all other operational programs, so the user would feel familiar with lists, check-buttons, window adjusters, etc.

The sentences illustrated in the concordance tool window can be categorized by frequency or alphabetically. These statistics are important for analysis of the context, like functional roles. The lines illustrate the user partner words of a search word.

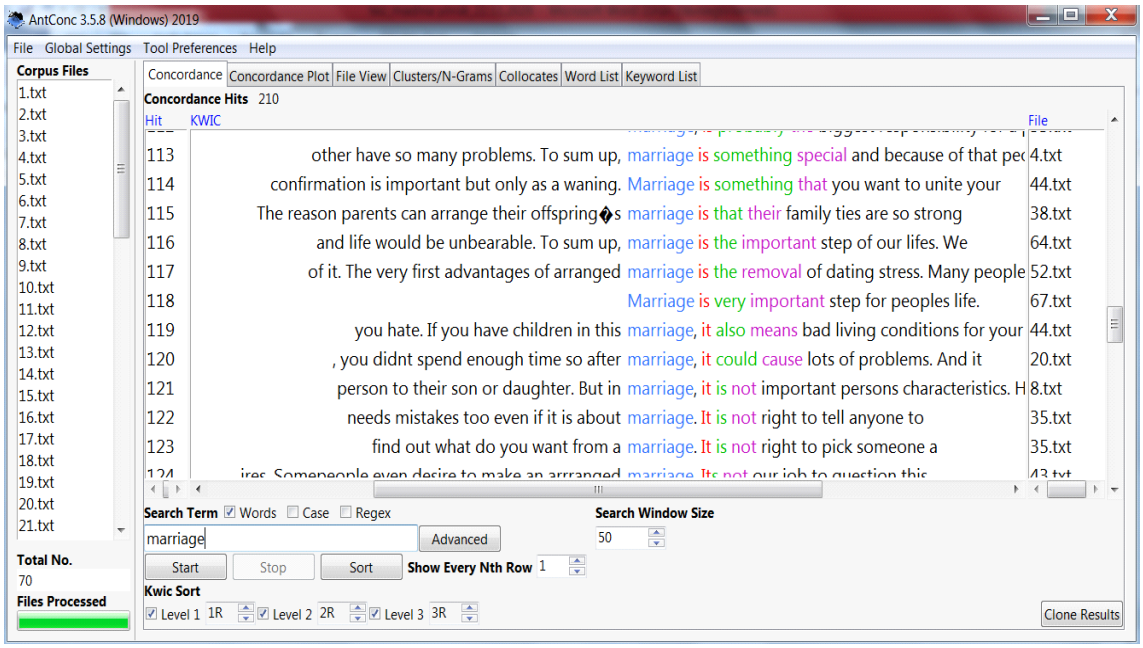

Figure 2.2. *KWIC Concordancer Tool*

This tool enables users to identify how a search word is located in a corpus. Concordance Search Term Plot Tool shows where exactly a search word is used, as illustrated in the figure below. Each line illustrates a file and files represent search term hits. This tool's purpose is to show the location of the search term, distribution options and the frequency number. If a user is searching for particular word or phrase, this tool might be an effective assistant.

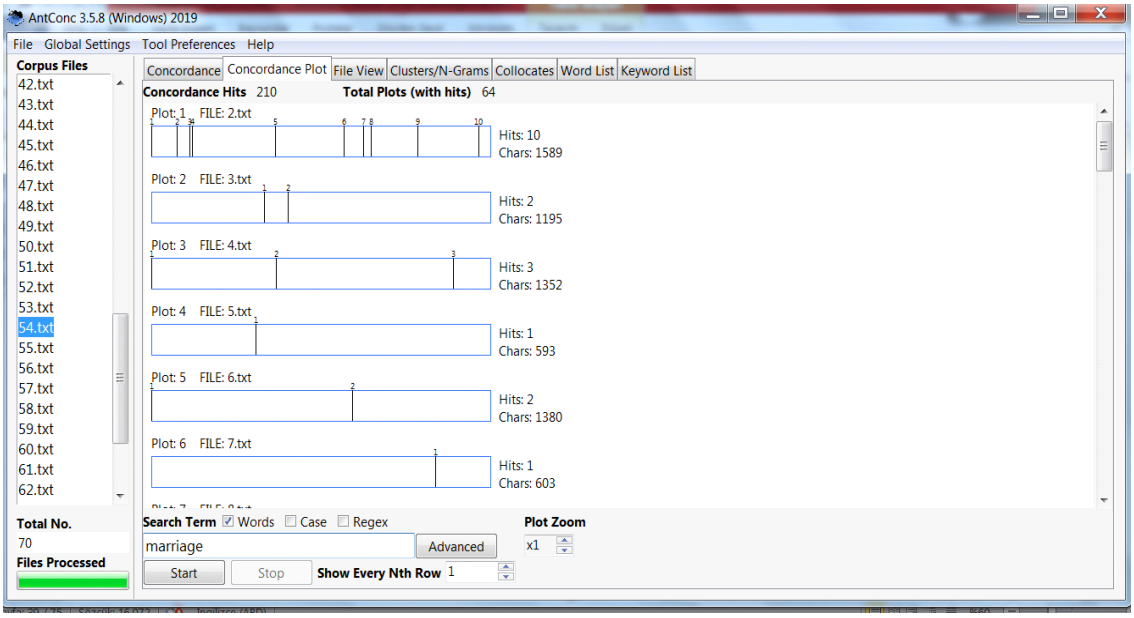

Figure 2.3. *Concordance Search Term Plot Tool*

As it was mentioned before, to see the location of a search term, a Concordancer Tool is used. The View Files Tool illustrates the search term in the initial text. Another advantage of this tool is that any phrase, word, regular expression or sub-string might be searched. The tool usually highlights hit words and enables to go to the next hit with keyboard shortcuts.

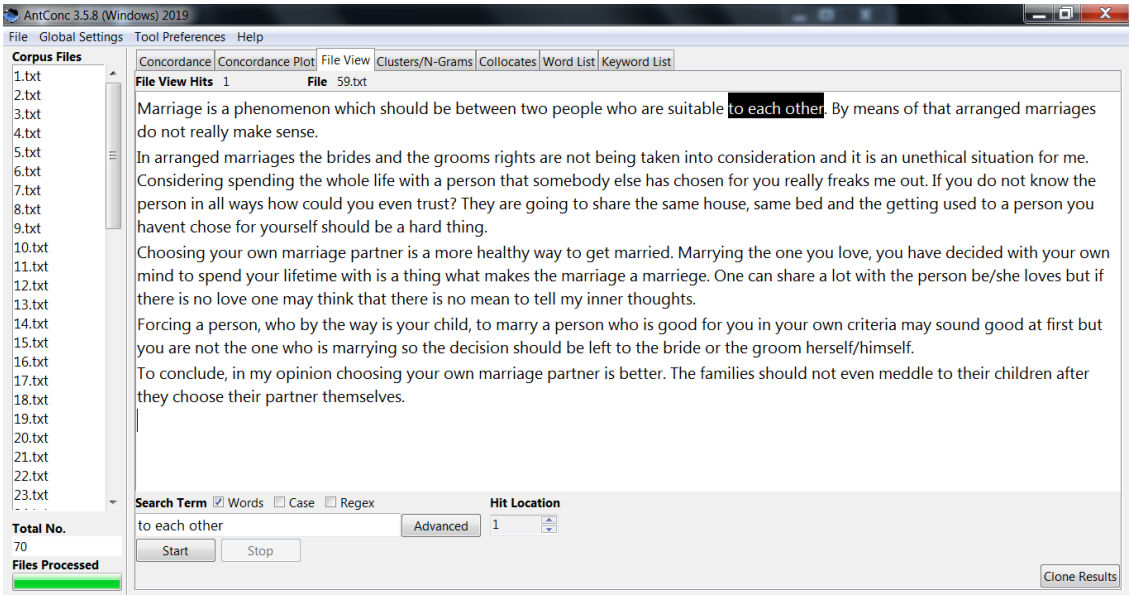

Figure 2.4. *View Files Tool*

In order to use a concordance program, it is necessary to upload all data to Antconc program and process it. The Keyword List Tool indicates incorrect area and shows different fields within the focus content. In addition, this tool is useful for highlighting similar words of target text or lemmas of words, as shown in Figure 2.5.

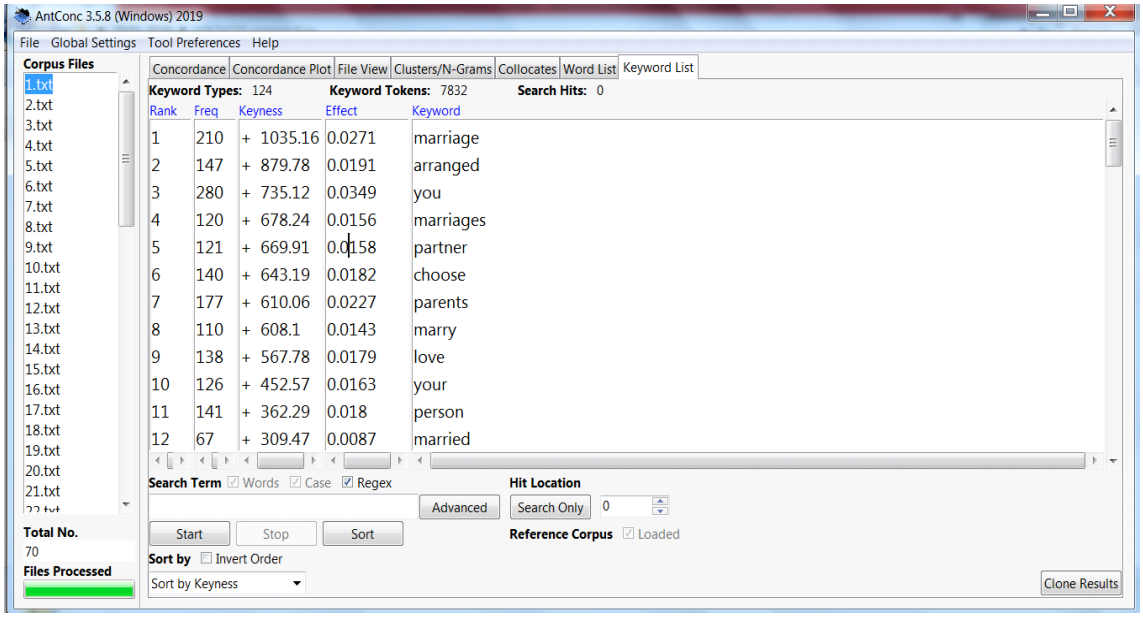

Figure 2.5. *Keyword List Tool*

The latest programs designed for word analysis have functions of indicating frequency order and list the words in alphabetical order. The word list function includes counting the number of words and it can divide them according to their root word. Moreover, a user can select a specific number of words to be processed and exclude the rest of the words. With regard to all mentioned above, a word plays a major role in a corpus. This program finds a word that repeats in the collected data and gives an overview for the researcher. The Keyword List tool is similar in Antconc and WordSmith programs.

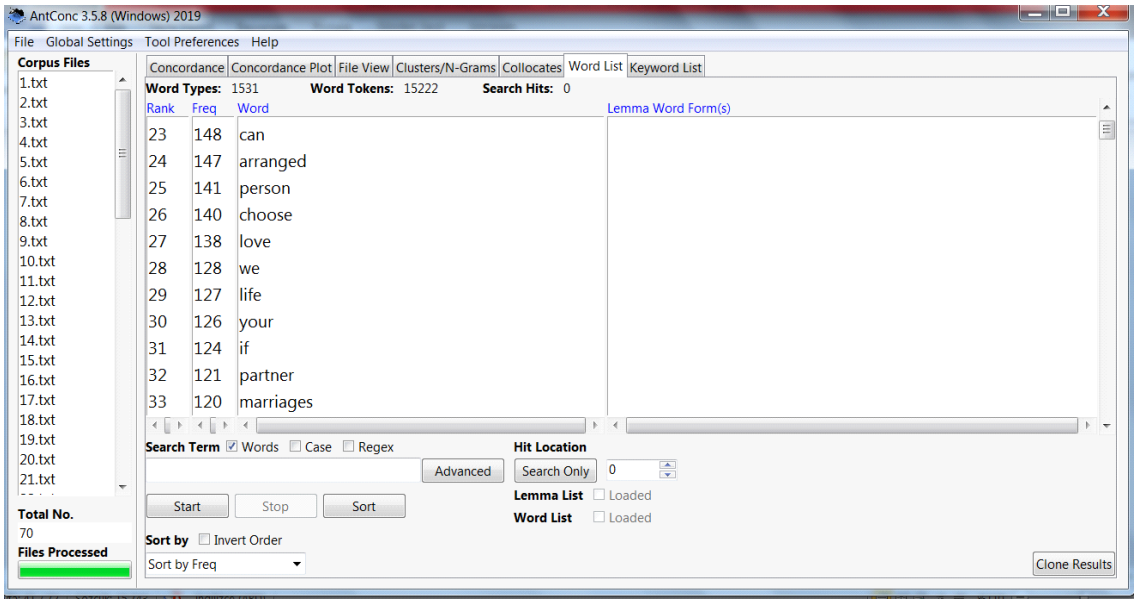

Figure 2.6. *Word List Tool*

The Word Clusters tool indicates multi-word units and sorts them by frequency in alphabetical order. As shown in the Figure 2.7, when we write a search word in the window of Word Clusters, a screen displays all occasions of the word being searched, with accompanying words before and after it. There is a line for a search word, a few search option boxes, cluster parameters, “save” box and progress report window.

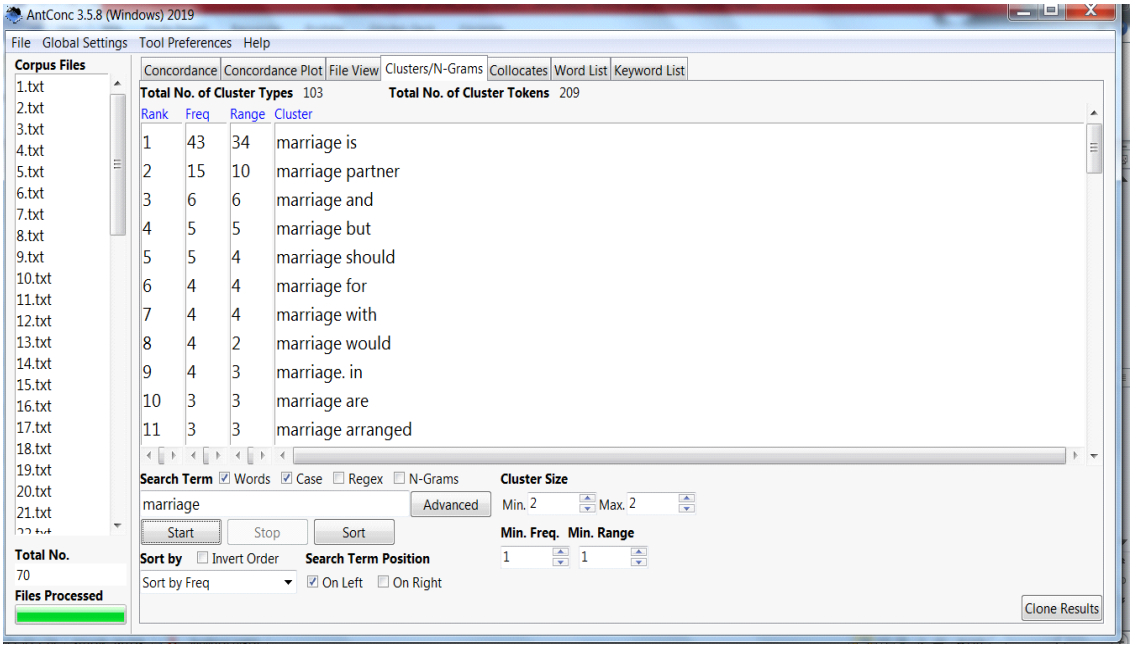

Figure 2.7. *Word Clusters*

There is no need to sort the data before downloading it to Antconc program. All data from a Microsoft Notepad document can be transferred to Antconc without sorting procedures. Anthony (2004) claimed that some other programs need a pre-processing of the words or amending the data before downloading it to a word processing program. Yet he admitted that an Antconc program is suitable only for a minor amount of data. Nevertheless, being free-of-cost makes this program competitive on the market.

Anthony (2004) stated that usually word analysis programs provide users with a table that is challenging to interpret and that the updated version of the Antconc program prepares a user-friendly result table. This new version of the program allowed users to easily copy and paste the results to other files, such as Microsoft Word and Microsoft Excel programs. A disadvantage of this program was that the encoding of data has to be in a specific HTML/XML format. Noguchi (2004) described the Antconc program as an effective, simple, lightweight and easy-to-use tool especially in classroom teaching. Even

though Antconc does not have some of the functions as other applications, which are not free of charge, it has vital instruments for data analysis, software license and a user-friendly interface. This program was upgraded 19 times; three of them were comprehensive upgrades (Anthony, 2004).

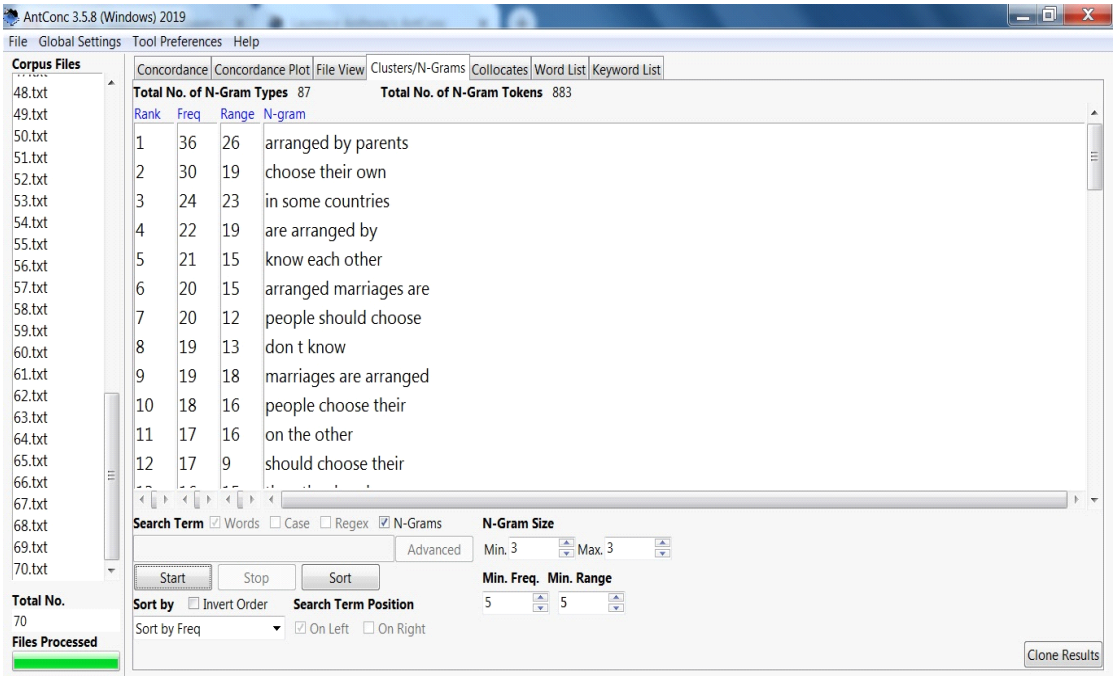

Figure 2.8. *Antconc program three-word bundles distribution*

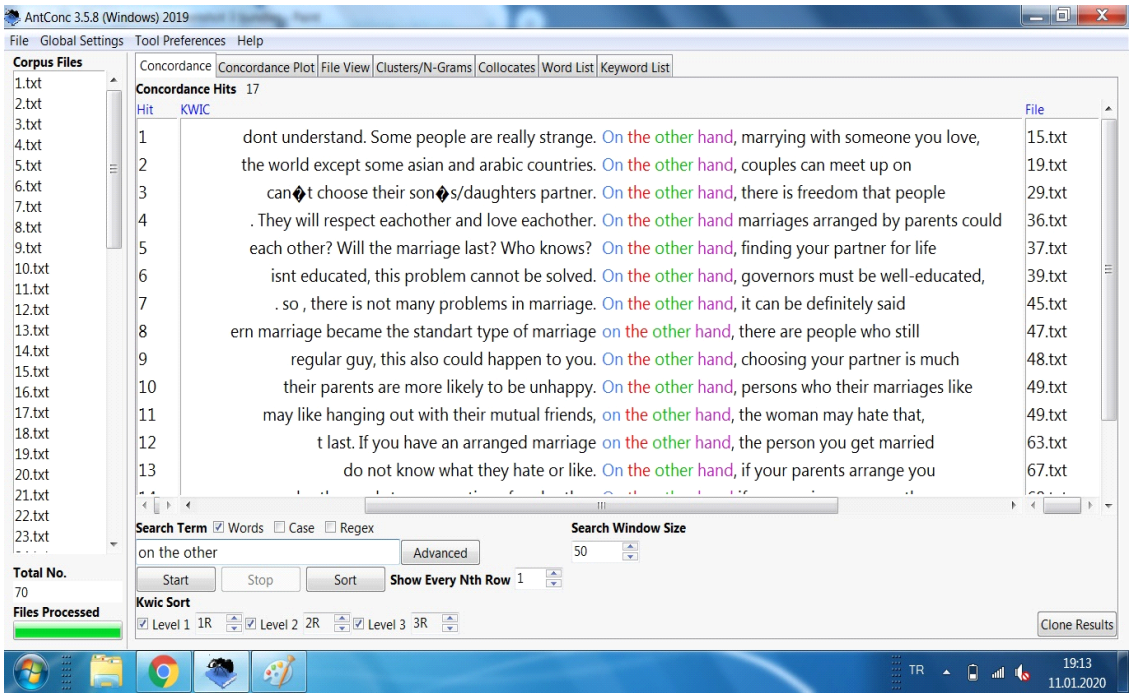

Figure 2.9. *Lexical bundle distribution*

3. METHODOLOGY

This part gives information about the whole process of the research, such as research corpus, data collection, structural classification and functional classification.

3.1. Research corpus

This study examined the writings of first year ELT students of three universities. While Turkish students were monolinguals, students from Kazakhstan were bilinguals with Russian language as a second mother language. One of these three state universities is located in Kazakhstan, while the other two are located in Turkey. Before elimination 250 argumentative essays were collected from three universities. After elimination 70 essays were processed from each university, thus making 210 essays in total. All 210 students were from the department of English language teaching. Two Turkish universities were located in Eskişehir city and one Kazakh University was in Almaty city. Kazakh university student were bilinguals since they all had two mother languages, Kazakh and Russian. The students from two Turkish universities were monolinguals.

Anadolu University is one of the leading universities in Turkey with around 2 million students. The entrance score for the faculty of education of Anadolu University is high and usually graduates of this faculty are successful in employment after graduation. Eskişehir Osmangazi University was divided from Anadolu University 30 years ago and the entrance score for the faculty of education is lower comparing to Anadolu University. Ablai Khan University, which is located in Kazakhstan, is in the top of the leading Kazakhstani universities in the field of linguistics, foreign languages and literature. The initial aim was to compare two leading universities of two countries in the field of foreign language teaching and compare two Turkish universities with the different university entrance score. Students in these countries have to pass a language exam in order to enroll in a four-year bachelor program. All students participating in this study were 1st year university students and they all successfully passed a language exam or completed one preparatory year at the university. Two hundred and fifty essays were collected and only 210 essays were analyzed after elimination (Appendix H and I). Some of essays were too short to be included in this study. An instructor from Ablai Khan University collected essays in one class hour. All students and university administration were informed of the research and the necessary permission was received. Similarly, two

Turkish university's management approved the data collection process and students were informed about the research. Essays from Kazakhstan were sent via email. All essays were handwritten and were entered onto computer for the analysis. The essays of two Turkish universities were written by Turkish students and the essays of Kazakh university were written by Kazakh and Russian students. Since this research is not focused on age and gender, this information is not presented in the research; however, it might be used in the future studies. Native speaker data was obtained from LOCNESS corpus. This corpus consists of essays written by British and American students. In order to compare 210 non-native essays with native essays, 217 argumentative essays were taken from LOCNESS corpus.

The data was collected from first year EFL students of three different universities on the following topic "In some countries, marriages are arranged by the parents. But in other cases, people choose their own marriage partner. Discuss both systems, and state which one do you think is better. Give reasons and examples for your answer". The students were asked to write an argumentative essay in a class hour. All handwritten essays were transferred to the electronic version for the analysis by Antconc program. Two hundred and fifty essays were examined from three universities. Biber's and Hyland's taxonomies were the main directions in this research, namely functional and structural classifications. The essays from Kazakhstan were received in pdf format by email. In the end, all lexical bundles were compared with the bundles of native speakers by means of LOCNESS corpus. In the structural analysis part, all lexical bundles were divided into four groups according to Biber's taxonomy, such as noun phrase bundles, verb phrase bundles, prepositional phrase bundles and eliminated bundles. After structural distribution the same lexical bundles were categorized according to Hyland's taxonomy, such as stance bundles, discourse organizers and referential bundles.

The argumentative essays in the LOCNESS corpus of native speakers were compared with the corpus of three universities. LOCNESS is a collection of native English essays designed by Belgium University which includes British and American students' work. It is a free corpus with 324,304 total number of words. Two hundred seventeen essays, namely argumentative essays of British and American students, were downloaded from LOCNESS corpus and analyzed through Antconc program.

3.2. EndcodeAnt Program

The analysis process was initiated by rewriting handwritten student papers to electronic format. Since the Antconc program processes only Notepad format data, essays which were in Microsoft word program were copied to the Notepad files. The handicap here was the impossibility of copying all student essays at once. It was needed to open a new Notepad file and transfer only one essay to it. Consequently, it took some time to transfer all rewritten data from Word document to a Notepad document. Each time, 70 essays of each university were downloaded to the Antconc program. Apart from this, Antconc program accepted only UTF-8 Unicode and as the data was in a different format, another program named EncodeAnt was downloaded. All essays were uploaded to EncodeAnt program and modified to UTF-8 Unicode as illustrated in figure 3.1.

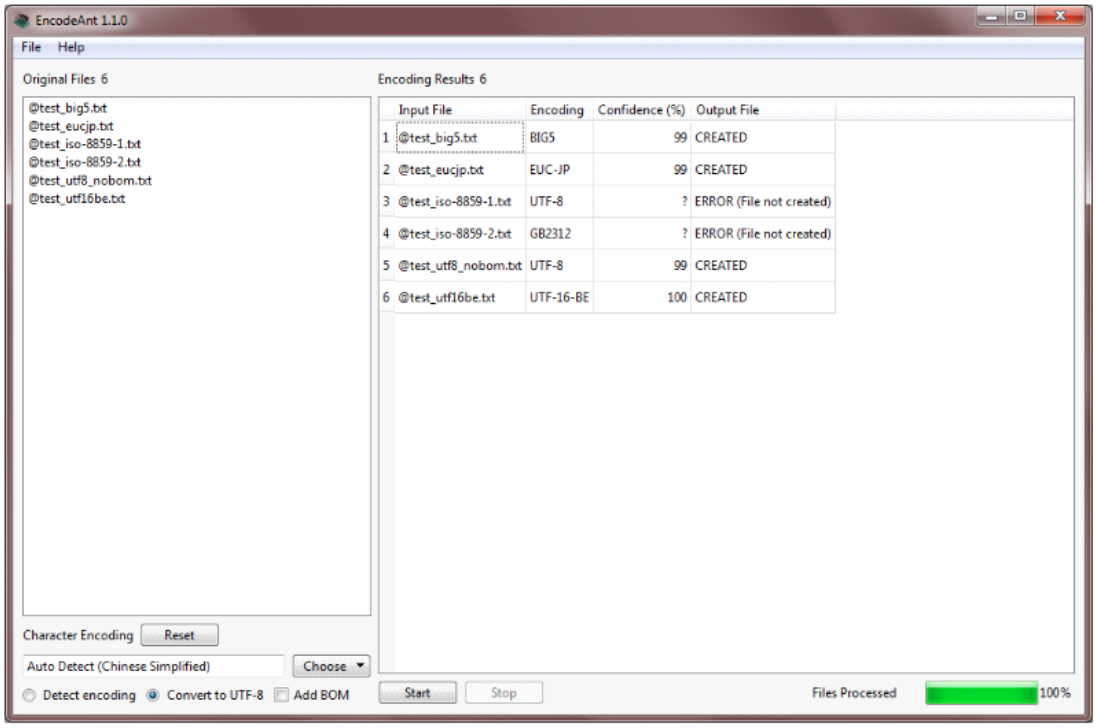

Figure 3.1. *EncodeAnt Program (https://www.laurenceanthony.net/)*

3.3. AntWordProfiler Program

In order to test the vocabulary level of students, Paul Nation's vocabulary test retrieved from https://www.wgtn.ac.nz/lals/about/staff/paul-nation, was implemented in all three groups. To do this test, all data was uploaded to AntWordProfiler program downloaded from https://www.laurenceanthony.net website as shown in figure 4.2.

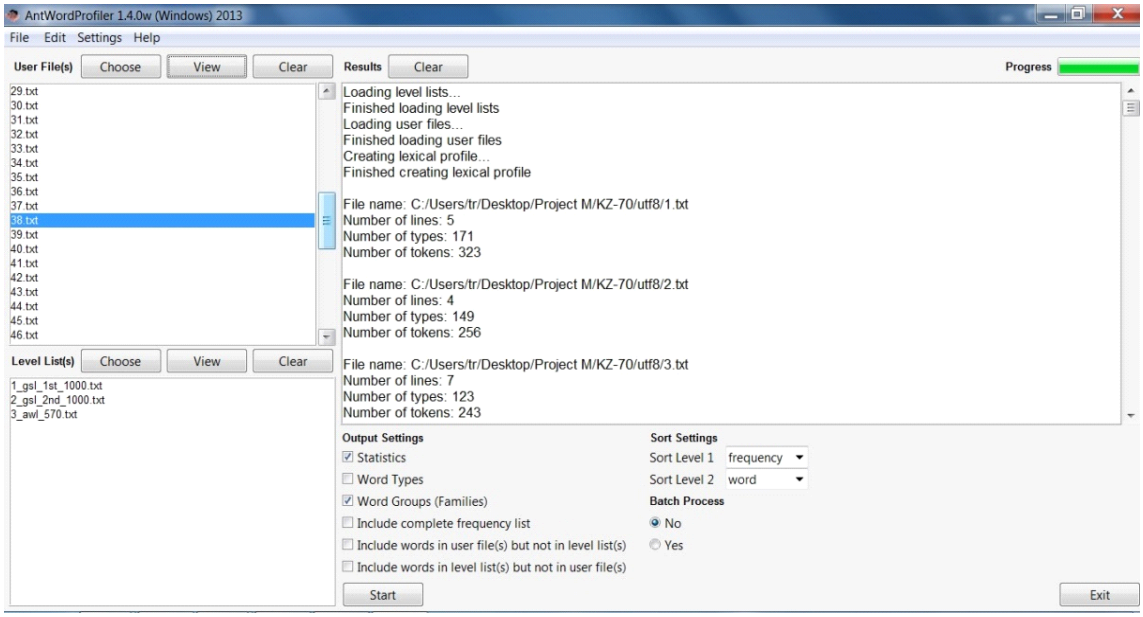

Figure 3.2. *AntWordProfiler Home Screen*

The results reported that the participant's level of English is within the basic vocabulary range as shown in Figure 3.3 and there is no big difference in the proficiency level of the students from three different universities.

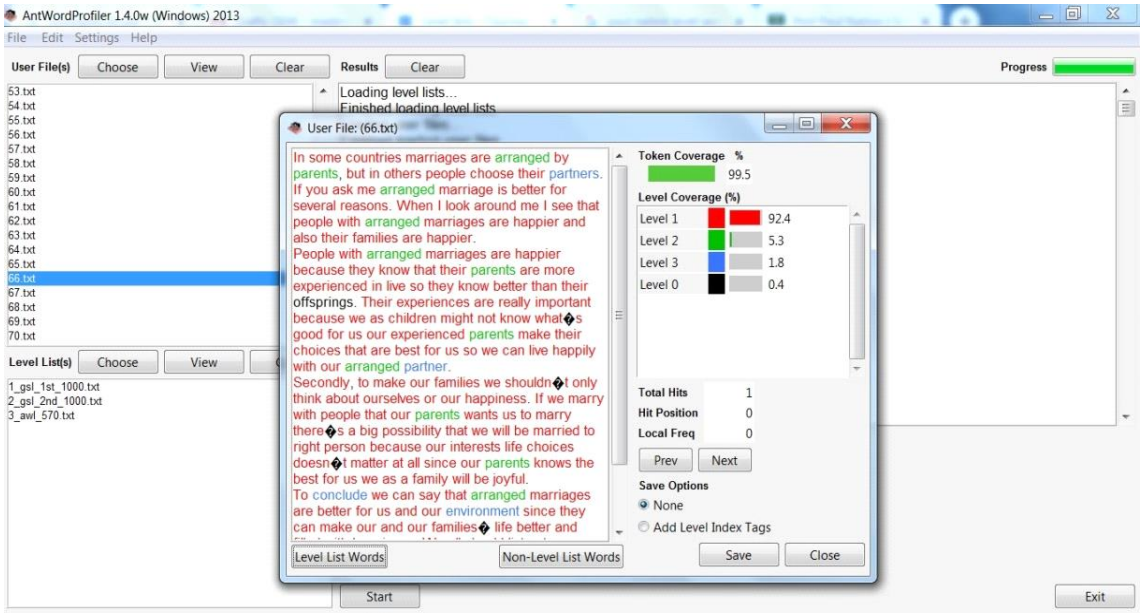

Figure 3.3. *AntWordProfiler Result Screen*

The AntWordProfiler program works in a similar way to Antconc. All data was uploaded through “Open files” tab, and after pressing “start” button the program presented a result window.

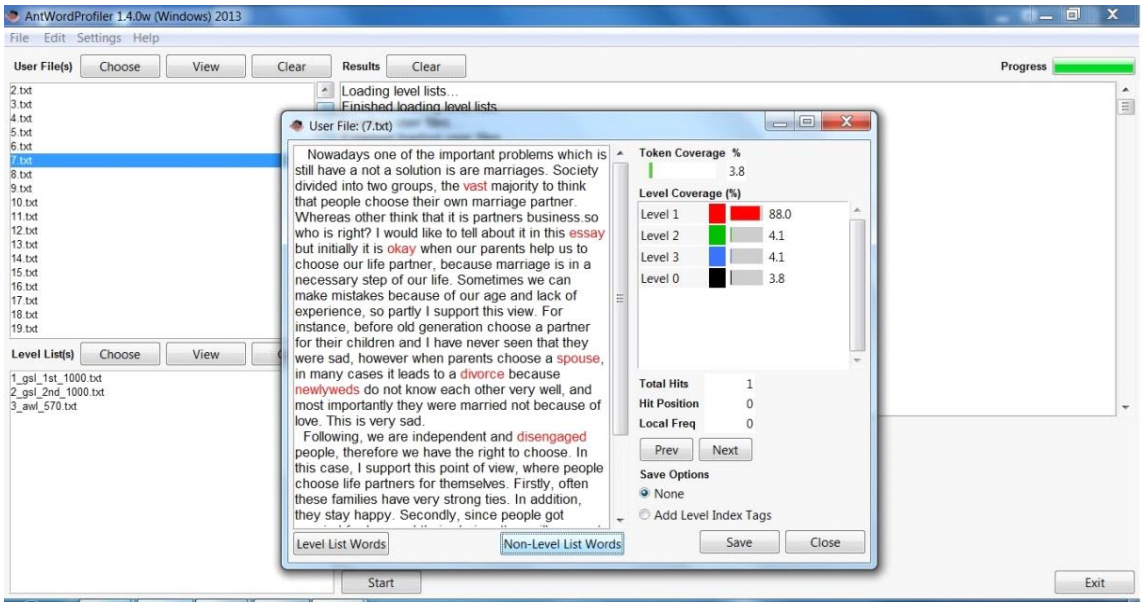

Figure 3.4. *AntWordProfiler Result Screen 2*

After selecting the length of bundles which were supposed to be investigated, the program gave the list of three-word bundles. After saving the results of all three university's students, the data was examined in Excel format. Columns of Excel helped to categorize lexical bundles in terms of structure and function. Antony Lawrence who created Antconc program has a few forums in the internet; these forums were helpful during the analysis part.

After uploading all the essays into the Antconc program, the following screen presented the results. In Figure 3.5., these results are displayed starting from the most frequent bundle, with the number of types and tokens.

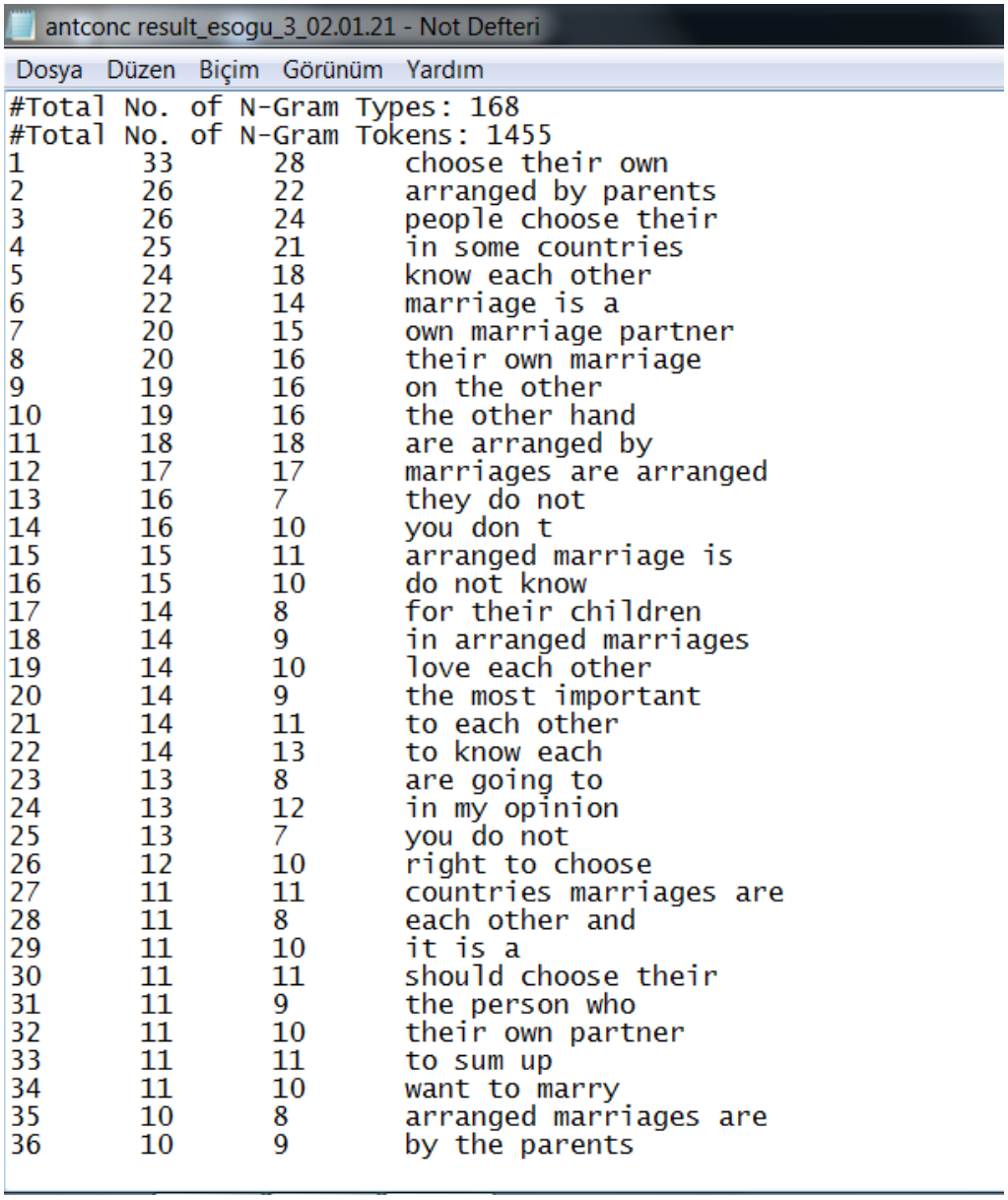

Figure 3.5. *Three-word lexical bundles distribution by Antconc program*

3.5. Structural Classification

As mentioned in the literature review part, one of the taxonomies utilized in this study is structural categorization by Biber. Here are twelve structural groups mentioned by Biber et al. in the Longman Grammar of Spoken and Written English (Biber et al., 1999).

Table 3.1. *Structural taxonomy (Biber et al.1999).*

Type	Example
Prepositional phrase with embedded *of*-phrase fragment	on the basis of, as a result of
Noun phrase with *of*-phrase fragment	the purpose of the, the end of the
Passive verb + prepositional phrase fragement	is related to the, are shown in the table
(Verb phrase +) that-clause fragment	that there is a, should be noted that
Noun phrase with other post-modifier fragments	the relationship between the, the extend to which
Adverbial clause fragment	if there is a, as shown in figure
Pronoun/noun + *be* + (…)	there was a significant, this is not the
(Verb/Adjective +) *to*-clause fragment	has been shown to, to be able to, are likely to be
Other preposition phrase (fragment)	on the other hand, in the present study
Anticipatory *it* + verb phrase/adjective phrase	it can be seen, it is possible to
Other expressions	than that of the, as well as the
Copula *be* + noun phrase/ adjective phrase	was no significant difference, is one of the

After five years, Biber et al. (2004) published more detailed structural categorization. This structural categorization has three main types. 1. Lexical bundles that incorporate *verb phrase* fragments, i.e. *you don't have to, it's going to be, I mean you know, is going to be, is based on the, are you going to, what do you think.* 2. Lexical bundles that incorporate dependent clause fragments, i.e. *I want you, what I want to, if you want to, to be able to* and *that there is a*. 3. Lexical bundles that incorporate *noun phrase* and *prepositional phrase* fragments, i.e. *one of the things, a little bit about, or something like that, at the end of* and *greater than or equal.*

Table 3.2. *Structural classification of lexical bundles (Biber et al. 2004).*

Types		Examples
Lexical bundles that incorporate verb phrase fragments	a. 1st/2nd person pronoun + VP fragment	*you don't have to, I'm not going to, well I don't know*
	b. 3rd person pronoun + VP fragment	*it's going to be, that's one of the, and this is a*
	c. discourse marker + VP fragment	*I mean you know, you know it was, I mean I don't*
	d. verb phrase (with non-passive verb)	*is going to be, is one of the, have a lot of, take a look at*
	e. verb phrase (with passive verb)	*is based on the, can be used to, shown in figure N*
	f. yes-no question fragments	*are you going to, do you want to, does that make sense*
	g. WH-question fragments	*what do you think, how many of you, what does that mean*
Lexical bundles that incorporate dependent clause fragments	a. 1st/2nd person pronoun + dependent clause fragment	*I want you to, I don't know if, I don't know why, you might want to*
	b. WH-clause fragments	*what I want to, what's going to happen, when we get to*
	c. If-clause fragments	*if you want to, if you have a, if we look at*
	d. to-clause fragments	*to be able to, to come up with, want to do is*
	e. that-clause fragment	*that there is, that I want to, that this is a*
Lexical bundles that incorporate noun phrase and prepositional phrase fragments	a. noun phrase with of-phrase fragment	*one of the things, the end of the, a little bit of*
	b. Noun phrase with other post-modifier fragment	*a little bit about, those of you who, the way in which*
	c. Other noun phrase expressions	*a little bit more, or something like that, and stuff like that*
	d. Prepositional phrase expressions	*of the things that, at the end of, at the same time*
	e. Comparative expressions	*as far as the, greater than or equal, as well as the*

Collected data was analyzed with regard to categorization by Biber et al. (1999). Three-word recurrent multi-word expressions, of the three university's students were divided into noun phase, verb phrase and prepositional phrase bundles.

3.5. Functional Classification

Hyland categorized lexical bundles into three broad groups with sub-categories (Hyland, 2008).

Table 3.3. *Functional taxonomy by Hyland (1999)*

Categories	Subcategories	Functions	Examples
Research-oriented	Location	Indicates time and place	at the same time
	Procedure	Indicates events, actions and methods	the use of the
	Quantification	Indicates quantities	a wide range of
	Description	Indicates property	an important role in
Text-oriented	Transitional signals	Establishes additive or contrastive links between elements	on the other hand
	Resultative signals	Makes inferential or causative relations between elements	the results of the
	Structuring signals	Organises streches of discourse or directs reader elsewhere in text	as shown in table
	Framing signals	Situates arguments by specifying limiting conditions	in the case of
Participant-oriented	Stance features	Conveys the writer's attitudes and evaluations	it is possible that
	Engagement features	Address readers directly	it should be noted

Statistical analysis of four corpora was performed by WordSmith program with the aim to identify whether the difference among corpora was statistically significant. Keyword function tool determined overused and underused bundles in the form of keyness calculations.

4. RESULTS AND DISCUSSION

4.1. Overall Results

The table below shows the total number of words in each corpus. The corpora of non-native speakers consist of 210 essays and 217 essays from native speakers' corpus were examined. Native speakers' data was downloaded from https://uclouvain.be/en/research-institutes/ilc/cecl/locness.html website. This corpus consists of British pupils' A level 60,209 word essays, British university students' 95,695 word essays and American university students' 168,400 word essays. Since 210 essays were written by three university's students, 217 essays were processed from LOCNESS native speaker corpus. To sum up, non-native speakers' corpora have 54,859 words and LOCNESS native speakers' corpus has 223,056 words. The total number of this research's corpora is 277,915 words.

Table 4.1. *Total number of words in all four corpora*

Institution	Number of Essays	Number of Words
Anadolu University	70	15231
Eskişehir Osmangazi University	70	18837
Ablai Khan University	70	20791
LOCNESS	217	223056
TOTAL		**277915**

Table 4.1. illustrates that LOCNESS corpus has the highest number of words. The difference between LOCNESS content and the other three university's student's writing content is clear. The three university's students wrote an essay on the same topic in the same period of time, while LOCNESS essays vary in terms of discipline and topic. Even though the number of essays of three universities is same, Ablai Khan University students have the highest number of words and Anadolu University's students show the least.

The table below shows the total number of lexical bundle types in four corpora.

Table 4.2. *Lexical bundle types in four corpora*

Corpus	Before	After	No of Texts
Anadolu University	202	33	70
Eskişehir Osmangazi University	168	32	70
Ablai Khan University	140	26	70
LOCNESS	222	55	217

As above table shows, some bundles were manually eliminated. The table shows the number of lexical bundles before and after elimination (Appendix A, B, C and D).

After examining the frequency cut off point in other research studies, it was agreed to use the frequency cut-off 25 times per 100,000 words. Since the size of all four corpora is different, the raw frequency was calculated as:

15,231 words ANADOLU corpus - 3.80 times (4 times)

18,837 words ESOGU corpus - 4.70 times (5 times)

20,791words Ablai Khan corpus - 5.19 times (5 times)

223,056 words LOCNESS corpus- 55.76 times (56 times).

These frequency calculations worked for three corpora except LOCNESS corpus. The frequency of LOCNESS corpus illustrated only 33 occurrences. After trying different combinations, it was decided to employ 20 times in three different texts version instead of 56 times in three different texts. The reason was that the frequency of this cut-off point results were the closest to the results of three other non-native corpora. Chen and Baker (2010) used 4 times in minimum 3 texts cut-off frequency, Grabowski and Jukneviciene (2018) preferred 4 times per 100,000 words and De Cock's (2004) cut-off point was 4 times per 100,000 words.

Table 4.3. *Most frequent three-word bundles of all four corpora*

Anadolu University		ESOGU		Ablai Khan University		LOCNESS	
on the other	17	on the other	19	**but in other**	24	the fact that	163
the other hand	16	the other hand	19	in my opinion	23	**in order to**	130
in my opinion	10	the most important	14	on the other	20	one of the	123
the rest of	10	**to each other**	14	one of the	20	there is no	94
a lot of	9	in my opinion	13	the other hand	20	due to the	82
but in other	9	**most of the**	10	first of all	15	the end of	82
most of the	9	one of the	10	point of view	15	**because of the**	80
first of all	8	**but in other**	9	a lot of	14	the idea of	77
as a result	7	there is no	9	the most important	14	there is a	77
there is a	7	is one of	8	this kind of	14	as well as	76
there is no	7	**the rest of**	8	purpose of this	12	end of the	70
if there is	6	a lot of	7	on their own	11	the use of	69
is an important	6	**in order to**	7	the purpose of	10	the number of	65
is the best	6	according to their	6	it is better	9	a lot of	61
one of the	6	at the end	6	there is no	9	part of the	59
a result of	5	for example in	6	is one of	8	on the other	58
because of the	5	is the best	6	my point of	8	at the end	56
for each other	5	is the most	6	there is a	8	the other hand	50
for the rest	5	**of the most**	6	**because of the**	7	**as a result**	45
is one of	5	point of view	6	**in order to**	7	out of the	45
it is also	5	rest of your	6	**of the most**	7	invention of the	44
there are a	5	some of them	6	is the most	6	the invention of	44
to each other	5	there are many	6	**to each other**	6	such as the	43
to have a	5	there are two	6	it is important	5	**the rest of**	43
it is important	4	there is a	6	the age of	5	the question of	41
rest of their	4	**as a result**	5	the choice of	5	all of the	40
rest of your	4	for this reason	5			is one of	39
the best for	4	in terms of	5			the case of	39
the end of	4	in this case	5			the amount of	38
the fact that	4	rest of their	5			the beginning of	38
the most important	4	the best for	5			**of the most**	37

Table 4.3. (Continue) *Most frequent three-word bundles of all four corpora*

the one you	4	the rate of	5	the majority of	37
this kind of	4			some of the	35
				most of the	34
				a part of	33
				a result of	33
				many of the	33
				a loss of	31
				according to the	30
				an example of	30
				in my opinion	29
				the effects of	28
				a sense of	25
				at the beginning	25
				beginning of the	25
				the concept of	25
				the lack of	25
				the loss of	25
				the most important	23
				the problem of	23
				majority of the	21
				rest of the	21
				a number of	20
				great deal of	20
				nature of the	20

After manual elimination 33 bundles of Anadolu University students, 32 bundles of ESOGU, 26 bundles of Ablai Khan University and 55 bundles of LOCNESS corpus were examined. Nine (9) common bundles were identified in the research and reference corpora. They are *on the other, the other hand, in my opinion, a lot of, there is a, there is no, is one of, the most important* and *one of the.* Common bundles in all four corpora are highlighted in brown in the table. Besides these nine common bundles there are six (6) bundles which come across in at least three corpora. They are *the rest of, but in other, most of the, as a result, because of the* and *to each other.* These six bundles are indicated in bold. These findings are similar to Yang's (2017) findings.

Anadolu University students and ESOGU students used most frequently the following two bundles: *on the other* and *the other hand*. Anadolu University students used *on the other* bundle 17 times and ESOGU students used 19 times. Bundle *the other hand* was used 16 times by Anadolu University students and 19 times by ESOGU students.

On the other *hand, couples can meet up on their own, in their everyday social life (Anadolu University-17).*

On the other *hand, love and respect is the base of a healthy family (ESOGU-19).*

On the other *hand, contractual marriage seems to be violent, restrictions of choice despite of cultural features of country, because according to the Asian mainland it is an unusual practice (Ablai Khan University-20).*

On the other *hand, it's easy to do things everyone expects of your gender (LOCNESS-58).*

Ablai Khan University students most frequently used *but in other* bundle, 24 times. This bundle was employed by only three non-native corpora.

In several countries, parents choose second half for their children, ***but in other*** *countries, newlyweds choose their spouse by themselves (Ablai Khan University-24).*

Second high frequency bundle shared by three non-native student groups is *to each other*. It was used five times by Anadolu University students, fourteen times by ESOGU students and six times by Ablai Khan University students.

Everything is bound ***to each other*** *(Anadolu University-5).*

They always are respectful ***to each other*** *(ESOGU-14).*

For example, if shapes of their heart are not perfect ***to each other****, this can cause divorce and cause psychological problem of their kids (Ablai Khan University-6).*

In the LOCNESS reference corpus the most frequently used bundle is *the fact that* which was used 163 times.

This is due to ***the fact that*** *human beings tend to become frustrated with those whom they like and love (LOCNESS-163).*

Second frequently used common bundle is *in order to* which is used 130 times and shared with non-native corpora. Such as, ESOGU students used it seven times and Ablai Khan University students used it seven times.

Couples need to try to have an equal partnership ***in order to*** *have a better marriage (LOCNESS-130).*

Another point to mention is that some bundles were employed only by non-native groups. For instance, *first of all, in terms of, the age of*, etc. These bundles were not found in the essays of native students. For instance, non-natives used *some of them* while natives preferred *the majority of*. Both of them express the same meaning; however, non-natives employed more primitive version.

Some of them *want to share life with person they love (ESOGU-6).*

The majority of *people just adopt this way of thinking without even thinking twice about its validity (LOCNESS-37).*

Another example is the bundle *there are many* which is employed only by non-natives while natives preferred to use *great deal of* bundle. These two examples indicated that both groups expressed the same meaning, although native speakers' expression is more professional.

Moreover, ***there are many*** *people who share that they want another partner on Facebook or write people who are younger than them (ESOGU-6).*

The school spent a ***great deal of*** *time training and finding jobs for graduates (LOCNESS-20).*

Two examples below similarly illustrated the same meaning, however were used by different L1 groups. *First of all* and *for example in* were used by non-natives and *at the beginning* and *such as the* by native students.

First of all*, the advantage of parents' choice is their wide range of life experience (Ablai Khan University-15).*

Children are under academic stress ***at the beginning*** *of the day (LOCNESS-25).*

For example, in *daily life, you cannot immediately trust someone you do not know well and share something (ESOGU-6).*

The government should do this by investing more money into public transport than into road improvements, ***such as the*** *widening of motorways (LOCNESS-43).*

In terms of bundle occurred five times in ESOGU student essays. Similarly, the frequency of *the age of* bundle is five times in Ablai Khan University student essays. *Point of view* bundle was used 21 times by non-native group. *In my opinion* bundle was also among frequent bundles. It occurred ten times in Anadolu University essays, 13 times in ESOGU essays, 23 times in KZ and 29 times in LOCNESS corpus.

4.2. Statistical Significance

Keyword function tool of the WordSmith program determined overused and underused bundles in the form of keyness calculations. Firstly, ANADOLU, ESOGU and Ablai Khan University students' word frequencies were compared to LOCNESS students. Secondly, LOCNESS data was compared to three non-native group frequency data. And lastly, two non-native groups were compared between each other.

Table 4.4. *ANADOLU and ESOGU bundles compared to LOCNESS corpus bundles.*

Corpus	Level	Lexical bundles	
		TO EACH OTHER	84,45
		BUT IN OTHER	84,45
		THE OTHER HAND	59,11
		IS THE BEST	56,30
		ON THE OTHER	55,47
		REST OF YOUR	46,92
		IN MY OPINION	42,36
Turkish	Overuse	REST OF THEIR	42,22
		ON THEIR OWN	42,22
		THE BEST FOR	42,22
		ACCORDING TO THEIR	37,53
		SOME PARTS OF	37,53
		THIS KIND OF	37,53
		THE MOST IMPORTANT	33,83
		SOME OF THEM	32,84
		AND YOU ARE	32,84
		FOR THE REST	32,84
	Underuse	---	

Table 4.5. *ANADOLU, ESOGU and Ablai Khan University bundles compared to LOCNESS corpus bundles*

Corpus	Level	Lexical bundles	
	Overuse	BUT IN OTHER	164,79
		TO EACH OTHER	94,16
		IN MY OPINION	89,19
		THIS KIND OF	86,31
		THE OTHER HAND	85,63
		ON THE OTHER	79,30
		IS THE BEST	62,77
		THE MOST IMPORTANT	59,22
		IT IS BETTER	54,92
Turkish+Kazakh		FIRST OF ALL	51,03
		THERE ARE SOME	51,00
		POINT OF VIEW	50,65
		PURPOSE OF THIS	47,08
		IS BETTER TO	43,15
		REST OF YOUR	43,15
		MY POINT OF	39,23

Table 4.5. (Continue) *ANADOLU, ESOGU and Ablai Khan University bundles compared to LOCNESS corpus bundles*

		BECAUSE OF THEIR	39,23
		SOME OF THEM	39,23
		IS THE MOST	29,82
	Underuse	---	

Table 4.6. *Ablai Khan University bundles compared to LOCNESS corpus bundles.*

Corpus	Level	Lexical bundles	
Kazakh	Overuse	BUT IN OTHER	139,61
		THIS KIND OF	81,43
		PURPOSE OF THIS	69,80
		IN MY OPINION	65,65
		POINT OF VIEW	58,43
		IT IS BETTER	52,35
		THERE ARE SOME	52,35
		MY POINT OF	46,53
		FIRST OF ALL	46,10
		THE OTHER HAND	38,19
		THE MOST IMPORTANT	35,79
		TO EACH OTHER	34,90
		IS BETTER TO	34,90
		ON THE OTHER	34,04
		THE PURPOSE OF	30,33
		BECAUSE OF THEIR	29,08
		BUT THERE ARE	29,08
		MY OPINION IT	29,08
		OF THEM IS	29,08
	Underuse	---	

A table above illustrated a variety of bundles used by four groups of students. These calculations confirm the previous studies where non-native speakers overused lexical bundles and native speakers underused (Appendix E, F, and G).

Apart from the analysis above, significance analysis of ANADOLU and ESOGU students compared to Ablai Khan University students did not show any occasion of underuse nor overuse. Similarly, keyness calculations for Ablai Khan University students

compared to ANADOLU and ESOGU students did not reported any occurrence of underuse and overuse.

Table 4.7. *LOCNESS corpus bundles compared to ANADOLU, ESOGU and Ablai Khan University bundles*

Corpus	Level	Lexical Bundles	
	Overuse	---	
LOCNESS	Underuse	IS THE MOST	-29,82
		POINT OF VIEW	-50,65
		FIRST OF ALL	-51,03
		THE MOST IMPORTANT	-59,22
		ON THE OTHER	-79,30
		THE OTHER HAND	-85,63
		IN MY OPINION	-89,19

The result in the above table shows that non-native speakers' lexical bundles are less varied. And overusing process is typical for non-native speakers of English.

One more comparison in addition to the table above was comparing LOCNESS corpus with only Kazakh university students, excluding Turkish students illustrated that all bundles are underused and there is no any occurrence of overused bundles. For instance, *the purpose of (-30,33), the other hand (-38,19), first of all (-46,10),* etc.

4.3. Structural Categorization

Table 4.8. showed the structural distribution of lexical bundles in percentage by Biber et al's classification (Biber et al., 1999). Overall numbers showed that noun phrases were prevalent in all four groups. While verb phrase bundles were the least used group of bundles. Anadolu University students used noun phrase bundles the most (65%), the percentage of prepositional phrase bundles (25%) and verb bundles (10%). ESOGU students mostly utilized noun phrase bundles (47%) and used verb phrase bundles (9%) the least. Ablai Khan University students, likewise LOCNESS students preferred noun phrase bundles (62%) in the first place, in the second place prepositional bundles (31%) and the last verb phrase bundles (7%). The calculations reported verb phrase bundles had the least occurrence in the essays of all four groups of students. The

results supported the findings of Bychkovska and Lee (2017) where native students preferred noun and prepositional phrase bundles. Esfandiari and Barbary's (2017) study similarly indicated that noun and prepositional phrase bundles were the most frequently used bundles. On the contrary, Chen and Baker's (2010) research yielded that the verb phrase bundles were the most frequently used group.

Table 4.8. *Structures of lexical bundles in four corpora in percentages*

Structure	ANADOLU	ESOGU	Ablai Khan	LOCNESS
Noun phrase bundles	65	47	62	75
Verb phrase bundles	10	9	7	2
Prepositional phrase bundles	25	44	31	23
Total	100	100	100	100

4.4. Functional Categorization

The structural and functional distribution of the recurrent word combinations was performed manually. Functional distribution is divided into following three types: 1. stance bundles 2. discourse organizers 3. referential expressions. According to Biber et al. (2004) stance expressions demonstrate "attitudes or assessments of certainty that frame some other positions".

Table 4.9. *Functional distribution of four corpora in percentages*

Corpus	Stance bundles	Discourse organizers	Referential expressions
ANADOLU	0	15	23
ESOGU	0	30	20
Ablai Khan	0	23	6
LOCNESS	0	32	42
	100	100	100

The above statistics showed that non-native speakers of Anadolu University and native speakers employed referential expressions most frequently. ESOGU and Ablai Khan University students preferred discourse organizers in the first place. Even though argumentative essays usually have the highest frequency of stance bundles, these four groups of students did not employed them. One assumption was that Anadolu University students still might not master subjects like topic introduction, topic elaboration and clarification, which are components of discourse organizers, that's was the reason of using discourse organizers less frequently than referential expressions.

"Since it is such a big deal, there are ***a lot of*** *debates over how it should be performed"* (ANADOLU).

In the example above Anadolu University students used quantity specification bundles.

"***On the other*** *hand, parents never wish evil to their child*" (Ablai Khan University).

"***The purpose of*** *this essay is to state that arranging marriage by parents is not the brilliant idea*" (Ablai Khan University).

"Perhaps ***the fact that*** *is a sport which almost entirely excludes women counts against it"* (LOCNESS).

"Also, ***a large number*** *of students don't live in the dorms"* (LOCNESS).

Above there is an example of topic elaboration bundle from discourse organizers used by Ablai Khan University students and framing attributes of referential expressions. And *a large number* bundle employed by LOCNESS students is an example of quantity specification of referential expressions.

5. CONCLUSION

5.1. Summary of the Study

The present study aimed to examine the use of lexical bundles in Turkish, Kazakhstani and English L1 speakers' writing, in the form of argumentative essay. It also served to unveil the lexical bundles shared by L1 Turkish, L1 Kazakh and L1 English students, in terms of structure and function. Particularly important was to find the types of lexical bundles that students share, some of which are distinctive only for native speakers and determine how they vary in terms of function and structure. Lexical bundles in argumentative essays of Anadolu University 1st year ELT students, Eskişehir Osmangazi University 1st year ELT students and Kazakh Ablai Khan University of International Relations and World Languages (Ablai Khan University) 1st year ELT students were analyzed and compared with The Louvain Corpus of Native English Essays (LOCNESS).

In the light of the findings of this study, they confirmed the general tendency of lexical bundles tradition that native speakers demonstrate to have less multi-word expressions than non-native speakers. It is important to mention that the level of English of non-native speakers was not advanced, since they were first year tertiary level students. Non-native speakers overused lexical bundles in comparison with native speakers. These calculations confirmed the previous studies where non-native speakers overused lexical bundles and native speakers underused.

The assumption was that non-native speakers exhibit a more restricted repertoire of n-grams and that this repertoire did not differ notably was not validated. For instance, varying wording were discovered with “there” structure and “this” structure, together with passive structures. A few points in the analysis appeared, which are worth mentioning. Comparison of research corpus with reference corpus revealed the influence of the topic on recurrent word combinations usage by all groups. For example, multi-word expressions of non-native speakers were topic-oriented, since all three groups of speakers produced an essay on a specific topic, while native speakers wrote essays in different genres.

In terms of structural categorization, overall numbers reported that noun phrases were prevalent in all four corpora. The calculations reported that verb phrase bundles had the least occurrence in the essays of all four groups of students. The results supported the findings of Bychkovska and Lee (2017) where native students preferred noun and

prepositional phrase bundles. Esfandiari and Barbary's (2017) study similarly indicated that noun and prepositional phrase bundles were the most frequently used bundles. On the contrary, Chen and Baker's (2010) research yielded that the verb phrase bundles were the most frequently used group. An explanation for these numbers might be the type of essay written. Yang (2017) stated that learners prefer more multi-word expressions in argumentative essays than in narrative essays. Even though all non-native students who took part in this research were sufficient EFL language users, their essays showed that they need to learn and broaden their academic vocabulary. The results also revealed that the number of lexical bundles used in the essays did not reflect the writers' academic competence, the right selection of the words and functional distribution were key points. Cortes (2020) reported that the students tried to use academic language in the essays, instead of conversational language (Cortes, 2020). For instance, native students almost did not use the word "I" which is often produced by three groups of non-native speakers in this research. It was expected to see the highest number of desire bundles, since they wrote argumentative essays. Anadolu University and native speakers employed referential expressions most frequently. ESOGU and Ablai Khan University students preferred discourse organizers in the first place. According to Chen and Baker (2010), the frequent use of discourse organizers is "a sign of immature writing". Ruan's (2016) study on Chinese undergraduate academic writing reported that discourse organizing bundles were the most prevalent bundles. Even though argumentative essays usually have the highest frequency of stance bundles, these four groups of students did not employed them.

Cortes (2004) and Schmidt (1990) encouraged instructors to show the students frequently used n-grams together with the functions in each academic text. In other words, help the learners to notice these lexical bundles and get familiar with their functions. The students of the three universities used colloquial expressions, as in the Shin's (2018) study.

The initial aim was to see the difference or similarities in lexical bundle use non-native students from two different countries and then compare them with native students. The different in two non-native groups was that Kazakhstani students were bilinguals and English language was their third language. The findings reported that non-native students from two countries had the same bundle usage in terms of structure. Both groups preferred noun phrase and prepositional phrase bundles. Functional classification revealed that ESOGU and Ablai Khan University students had similarities and Anadolu University

students had different performance. The reason can be that Anadolu University entrance criteria is higher comparing to other two universities. Native speakers' functional and structural analysis indicated almost similar results with non-native students. However, statistical analysis revealed that non-native speakers overused frequently recurring word sequences in comparison with native speakers.

5.2. Implications and Suggestions for Further Study

Lexical bundles have been found to be employed more by non-native speakers rather than native ones. Due to the globalization and availability of programs in improving English language skills with the help of modern technologies, students have access to a large amount of data and they strive to show as much as possible the presence of advanced knowledge of the language. For non-native instructors seeking to improve their students' writing skills it is recommended to pay attention to recurrent word combinations and the importance of not overusing them "to make students more aware of these structures and reach native like performance" (Muşlu, 2018). Future research may include rating argumentative essays of L1 and L2 speakers by native English teachers and distinguish their performance difference. This will result to evaluating each work and determining the average number of multi-word combinations allowed in one academic work to achieve the maximum score. Moreover, sort them according to the structural differences mentioned in Nekrasova's (2009) study on the perception of some lexical bundles more by L1 speakers rather than L2 speakers. Academic papers with greater number of bundles cannot be defined as good one or vice versa. Next step can be creating a concept of bundle usage and implementing it in English language textbooks in different levels. This mentioned guide may be useful for language instructors who may not be native English speakers. Byrd (2010) stated that "Focusing just on lexical bundles would not give students access to the full range of formulaic, multi-word units that are regularly used in academic writing", but will attract their attention on these important aspects of the language. One should note that language tend to vary depending on multiculturalism: colloquial speech may change and new words may be occurred. The expressions that are clear to the recipient (an ELT teacher and a student with the same cultural background), may not be acceptable to use in the academic writings (essay, article, etc.). Yoon and Choi (2015) research reported that "overall lack of register awareness in English argumentative writing" of Korean students which is also seen in Turkish and Kazakhstani university

students' works and demonstrates the general relevance of the topic. Regardless of the level of language proficiency, L2 students try to learn the basics of academic literacy. And bundles are an integral part of it, understanding, being aware and able to use them can significantly increase the level of written statements.

Lexical bundles teaching might be combined with academic writing subjects at schools. Lexical bundles topic could be added to the course syllabus and students may learn this term and read empirical studies related to lexical bundles. There is a possibility that students after getting familiar with overused bundles by non-natives situation, could improve their lexical bundle usage as EFL learners. Students may categorize bundles in the class in terms of structure and functions and consequently be familiar with all lexical bundle process. Such categorization task could be an effective practice from morphological aspect. As Öztürk (2014) mentioned, students can do some exercises in the class and find lexical bundles employed in specific area, for instance as in Allen's (2009) study on soft and hard sciences.

There are not many studies on different subject areas like Allen's (2009). Researchers may focus on this field and identify lexical bundles in medicine, informatics, marketing, etc. In addition, different L1 speakers could be involved in lexical bundle research and it would be easier to identify whether it is transfer from L1 as mentioned in the research of Paquot (2013). The last suggestion is that all these L1 speakers' corpora could be collected in one large corpus of lexical bundles expressed by non-native speakers of English.

REFERENCES

Adel, A., & Erman, B. (2012). Recurrent word combinations in academic writing by native and non-native speakers of English: A lexical bundles approach. *English for Specific Purposes, 31*(2), 81-92.

Allen, D. (2009). Lexical bundles in learner writing: An analysis of formulaic language in the ALESS Learner Corpus. *Komaba Journal of English Education, 1*(4), 105-127.

Anthony, L. (2004). AntConc: A learner and classroom friendly, multi-platform corpus analysis toolkit. Proceedings of IWLeL, 7-13.

Anthony, L. (2019). AntConc (Version 3.5.8) [Computer Software]. Tokyo, Japan: Waseda University. Available from https://www.laurenceanthony.net/software

Bal, B. (2010). *Analysis of Four-word Lexical Bundles in Published Research Articles Written by Turkish Scholars.* Unpublished MA Thesis. Georgia State University, U.S.A

Biber, D., Douglas, B., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use.* Cambridge: Cambridge University Press.

Biber, D., & Conrad, S. (1999). Lexical bundles in conversation and academic prose. *Language and Computers, 26*, 181-190.

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (2000). Longman grammar of spoken and written English. *The Electronic Journal for English as a Second Language, 4*(3), 7-21.

Biber, D., Conrad, S., & Cortes, V. (2004). If you look at…: Lexical bundles in university teaching and textbooks. *Applied linguistics, 25*(3), 371-405.

Biber, D., & Barbieri, F. (2007). Lexical bundles in university spoken and written registers. *English for specific purposes, 26*(3), 263-286.

Biber, D., Kim, Y. & Tracy-Ventura, N. (2010). A corpus-driven approach to comparative phraseology: lexical bundles in English, Spanish, and Korean. *Japanese/Korean Linguistics. 17*, 75-94.

Bychkovska T., & Lee, J. J. (2017). At the same time: Lexical bundles in L1 and L2 university student argumentative writing. *Journal of English for Academic Purposes, 30*, 38-52.

Byrd, P., & Coxhead, A. (2010). On the other hand: Lexical bundles in academic writing and in the teaching of EAP. *University of Sydney Papers in TESOL*, *5*(5), 31-64.

Chen, Y., & Baker, P. (2010). Lexical Bundles in L1 and L2 Academic Writing. *Language Learning & Technology, 14*(2), 30–49.

Conrad, S. and D. Biber. (2004). The frequency and use of lexical bundles in conversation and academic prose. *Lexicographica: International Annual for Lexicography*, *20*, 56-71.

Cortes, V. (2002). Lexical bundles in Freshman composition. In R. Reppen, S. M. Fitzmaurice & D. Biber (Eds.), Using corpora to explore linguistic variation (pp. 131–145). Amsterdam: John Benjamins Publishing Company.

Cortes, V. (2004). Lexical bundles in published and student disciplinary writing: Examples from history and biology. *English for Specific Purposes*, *23*, 397–423.

Cortes, V. (2013). The purpose of this study is to: Connecting lexical bundles and moves in research article introductions. *Journal of English for academic purposes, 12*(1), 33-43.

Coxhead, A., & Byrd, P. (2007). Preparing writing teachers to teach the vocabulary and grammar of academic prose. *Journal of second language writing, 16*(3), 129-147.

De Cock, S. (2004). Preferred sequences of words in NS and NNS speech. *Belgian Journal of English Language and Literatures (BELL), 2*(1), 225-246.

Dontcheva-Navratilova, Olga. (2012). Lexical Bundles in Academic Texts by Non-native Speakers. *Brno Studies in English, 38*, 37-58. 10.5817/BSE2012-2-3.

Erman, B. (2009). Formulaic language from a learner perspective: What the learner needs to know. In B. Corrigan, H. Quali, E. Moravcsik, & K. Wheatley (Eds.), Formulaic language (pp. 27–50). Amsterdam: John Benjamins.

Esfandiari, R., & Barbary, F. (2017). A contrastive corpus-driven study of lexical bundles between English writers and Persian writers in psychology research articles. *Journal of English for Academic Purposes, 29*, 21-42.

Granger, S. (1998). The computer learner corpus: A versatile new source of data for SLA research. In Granger, S. (ed.) Learner English on Computer. Addison Wesley Longman: London & New York, 3-18.

Gezegin-Bal, B. (2019). Lexical bundles in published research articles: A corpus-based study. *Journal of Language and Linguistic Studies, 15*(2), 520-534.

Juknevičienė, R., & Grabowski, Ł. (2018). Comparing formulaicity of learner writing through phrase-frames: A corpus-driven study of Lithuanian and Polish EFL student writing. *Research in Language*, *16*(3), 303-323.

Gray, B., and D. Biber. 2013. Lexical Frames in Academic Prose and Conversation. *International Journal of Corpus Linguistics 18.*109–135.

Güngör, F., & Uysal, H. H. (2016). A Comparative Analysis of Lexical Bundles Used by Native and Non-native Scholars. *English Language Teaching*, *9*(6), 176-188.

Howarth, P. (1998). Phraseology and second language proficiency. *Applied Linguistics*, *19*(1), 24–44.

Hyland, K. (2008). As can be seen: Lexical bundles and disciplinary variation. *English for specific purposes, 27(1), 4-21.*

Hyland, K. (2012). Bundles in academic discourse. *Annual review of applied linguistics, 32,* 150-169.

Juknevičienė, R. (2009). Lexical bundles in learner language: Lithuanian learners vs. native speakers. *Kalbotyra, 61,* 61-72.

Karabacak, E., & Qin, J. (2013). Comparison of lexical bundles used by Turkish, Chinese, and American university students. *Procedia-Social and Behavioral Sciences, 70,* 622-628.

Kashiha, H., & Heng, C. S. (2013). An exploration of lexical bundles in academic lectures: examples from hard and soft sciences. *The Journal of AsiaTEFL, 10*(4).

Kumaravadivelu, B. (2001). Toward a postmethod pedagogy. *TESOL quarterly, 35*(4), 537-560.

Lee, Y. E., Yoo, I. W., & Shin, Y. K. (2020). The use of English prepositions in lexical bundles in essays written by Korean university students. *Journal of English for Academic Purposes, 45,* 100848.

Lewis, M. (2008). *The idiom principle in L2 English: Assessing elusive formulaic sequences as indicators of idiomaticity, fluency, and proficiency* (Doctoral dissertation, Engelska institutionen). Stockholm University, Faculty of Humanities, Department of English.

Liu, D. (2012). The most frequently-used multi-word constructions in academic written English: A multi-corpus study. *English for Specific Purposes, 31*(1), 25-35.

Liu, C. Y., & Chen, H. J. H. (2020). Analyzing the functions of lexical bundles in undergraduate academic lectures for pedagogical use. *English for Specific Purposes, 58,* 122-137.

Liu, C. Y., & Chen, H. J. H. (2020). Functional variation of lexical bundles in academic lectures and TED talks. *Register Studies, 2*(2), 176-208.

Lonfils, C., & Vanparys, J. (2001). How to design user-friendly CALL interfaces. *Computer Assisted Language Learning, 14*(5), 405-417.

Lu, X., & Deng, J. (2019). With the rapid development: A contrastive analysis of lexical bundles in dissertation abstracts by Chinese and L1 English doctoral students. *Journal of English for Academic Purposes, 39,* 21-36.

Mahlberg, M. (2007). Corpus stylistics: bridging the gap between linguistic and literary studies. Text, discourse and corpora: *Theory and analysis, 8,* 219-246.

McCarthy, M. 2004. Touchstone: From Corpus to Course Book. Cambridge: Cambridge University Press

McEnery, T., & Hardie, A. (2011). What is corpus linguistics? In Corpus Linguistics: Method, Theory and Practice (Cambridge Textbooks in Linguistics, pp. 1-24). Cambridge: Cambridge University Press.

Muşlu, M. (2018). Use of stance lexical bundles by Turkish and Japanese EFL learners and native English speakers in academic writing. *Gaziantep Üniversitesi Sosyal Bilimler Dergisi, 17*(4), 1319-1336.

Nekrasova, T. M. (2009). English L1 and L2 speakers' knowledge of lexical bundles. *Language learning, 59*(3), 647-686.

Noguchi, J. (2004). A genre analysis and mini-corpora approach to support professional writing by non-native English speakers. *English Corpus Studies, 11,* 101-110.

Öztürk, Y. (2014). Lexical bundle use of Turkish and native English writers: A corpus-based study. *Unpublished MA dissertation.* Anadolu University, Eskişehir, Turkey.

Öztürk, Y., & Köse, G. D. (2016). Turkish and native English academic writers‟ use of lexical bundles. *Journal of Language and Linguistic Studies, 12*(1), 149-165.

Paquot, M. (2013). Lexical bundles and L1 transfer effects. *International Journal of Corpus Linguistics, 18*(3), 391-417.

Pan, F., Reppen, R., & Biber, D. (2016). Comparing patterns of L1 versus L2 English academic professionals: Lexical bundles in Telecommunications research journals. *Journal of English for Academic Purposes, 21,* 60-71.

Pan, F., & Liu, C. (2019). Comparing L1-L2 differences in lexical bundles in student and expert writing. *Southern African Linguistics and Applied Language Studies, 37*(2), 142-157.

Pang, W. (2010). Lexical bundles and the construction of an academic voice: A pedagogical perspective. *Asian EFL Journal, 47,* 1–13.

Qin, J. (2013). Use of formulaic bundles by non-native English graduate writers and published authors in applied linguistics. *System, 42,* 220-231.

Rica-Peromingo, J. P. (2009). The use of lexical bundles in the written production of Spanish EFL university students. Proceedings. Riga: University of Latvia Publishing.

Römer, U., & Arbor, A. (2009). English in academia: Does nativeness matter. Anglistik: *International Journal of English Studies, 20*(2), 89-100.

Ruan, Z. (2017). Lexical Bundles in Chinese Undergraduate Academic Writing at an English Medium University. *RELC Journal, 48*(3), 327–340.

Salazar, L., & Joy, D. (2011). *Lexical bundles in scientific English: A corpus-based study of native and non-native writing* (Doctoral dissertation, Universitat de Barcelona).

Salazar, D. (2014). Lexical bundles in native and non-native scientific writing: *Applying a corpus-based study to language teaching Vol. 65.* John Benjamins Publishing Company.

Schmidt, R. W. (1990). The role of consciousness in second language learning. *Applied linguistics, 11*(2), 129-158.

Shin, Y. K. (2018). The Construction of English Lexical Bundles in Context by Native and Nonnative Freshman University Students. *English Teaching, 73*(3), 115-139.

Shin, Y. K. (2019). Do native writers always have a head start over non-native writers? The use of lexical bundles in college students' essays. *Journal of English for Academic Purposes, 40,* 1-14.

Staples, S., Egbert, J., Biber, D. & McClair, A. (2013). Formulaic sequences and EAP writing development: Lexical bundles in the TOEFL iBT writing section. *Journal of English for Academic Purposes, 12,* 214-225.

Suethanapornkul, S. (2009). A story in four words: An analysis of lexical bundles in learners' writing placement test in ELIPT corpus.

Sun, Y. C., & Wang, L. Y. (2003). Concordancers in the EFL classroom: Cognitive approaches and collocation difficulty. *Computer Assisted Language Learning*, *16*(1), 83-94.

Tenuta, A. M., Oliveira, A. L. A., & Orfanó, B. M. (2012). How Brazilian learners express modality in their writing: A corpus-based study on lexical bundles. *Intercâmbio. Revista do Programa de Estudos Pós-Graduados em Linguística Aplicada e Estudos da Linguagem. ISSN 2237-759X*, *26.*

Uysal, H. H. (2012). Argumentation across L1 and L2 Writing: Exploring Cultural Influences and Transfer Issues. *Vigo International Journal of Applied Linguistics*, *9*, 133-159.

Wei, Y., & Lei, L. (2011). Lexical bundles in the academic writing of advanced Chinese EFL learners. *RELC journal*, *42*(2), 155-166.

Yang, Y. (2017). Lexical bundles in argumentative and narrative writings by Chinese EFL learners. *International Journal of English Linguistics, 7*(3), 58-69..

Zamel, V. (1998). Questioning academic discourse. *Negotiating academic literacies: Teaching and learning across languages and cultures*, 187-197.

http-1: https://www.wgtn.ac.nz/lals/about/staff/paul-nation

(Erişim tarihi: 10.03.2020)

APPENDICES

Appendix A – Three-word bundles of Ablai Khan University students

antconc result_KZ_3_02.01.21 - Not Defteri

Dosya Düzen Biçim Görünüm Yardım

#Total No. of N-Gram Types: 140
#Total No. of N-Gram Tokens: 1406

1	41	29	own marriage partner
2	39	29	choose their own
3	35	29	people choose their
4	33	28	their own marriage
5	31	26	by the parents
6	31	26	in some countries
7	26	22	arranged by the
8	24	24	but in other
9	23	19	for their children
10	23	21	in my opinion
11	21	18	are arranged by
12	21	21	parents but in
13	20	19	i want to
14	20	17	on the other
15	20	16	one of the
16	20	18	the other hand
17	19	18	marriages are arranged
18	18	18	in other cases
19	18	17	in this essay
20	17	17	cases people choose
21	17	17	the parents but
22	16	16	other cases people
23	15	15	first of all
24	15	14	point of view
25	15	14	this essay i
26	15	15	to sum up
27	14	12	a lot of
28	14	14	the most important
29	14	13	this kind of
30	13	13	countries marriages are
31	13	11	partner for their
32	13	13	some countries marriages
33	12	12	of this essay
34	12	12	purpose of this
35	12	9	to choose their
36	11	8	on their own
37	11	10	would like to
38	10	10	essay i will
39	10	9	i would like
40	10	10	in conclusion i
41	10	10	the purpose of
42	10	10	this essay is
43	9	9	both sides and
44	9	7	choose marriage partner
45	9	9	essay is to
46	9	7	it is better
47	9	8	it is not
48	9	8	it will be
49	9	9	love each other
50	9	8	marriage partner for
51	9	8	that it is
52	9	9	there are some
53	9	9	there is no
54	9	9	to say that
55	9	9	will discuss both
56	8	8	by their parents
57	8	4	choose the partner
58	8	8	discuss both sides
59	8	7	each other and
60	8	8	i think that
61	8	8	is one of
62	8	8	my point of
63	8	8	of their parents
64	8	7	there is a
65	8	8	want to say
66	7	4	advantages and disadvantages
67	7	7	because it is
68	7	7	because of the
69	7	7	i will discuss
70	7	7	in order to
71	7	7	in other countries
72	7	6	in the past
73	7	6	marriage is a
74	7	7	of the most
75	7	7	should choose their
76	7	7	that people should
77	7	6	to get married
78	7	6	will live with
79	7	5	your marriage partner
80	6	6	are going to
81	6	5	as we know
82	6	5	be better if
83	6	5	by looking at
84	6	6	conclusion i want
85	6	6	for their child
86	6	6	have to be
87	6	6	in some cases
88	6	5	in the life
89	6	5	is better to
90	6	5	is the most
91	6	6	is to discuss
92	6	5	know each other
93	6	6	marriage partner is
94	6	6	of our life
95	6	5	of your life
96	6	6	one of them
97	6	6	parents choose marriage
98	6	6	people believe that
99	6	6	people should choose
100	6	5	right to choose
101	6	5	they want to

antconc result_KZ_3_02.01.21 - Not Defteri

Dosya Düzen Biçim Görünüm Yardım

102	6	6	think that it
103	6	6	to each other
104	5	4	and now they
105	5	5	because of their
106	5	4	but there are
107	5	5	can say that
108	5	4	choice of a
109	5	5	choose their partner
110	5	5	choose their partners
111	5	5	choosing a profession
112	5	5	choosing marriage partner
113	5	5	essay will discuss
114	5	5	for example in
115	5	4	going to marry
116	5	4	his or her
117	5	4	i do not
118	5	4	in our life
119	5	4	in the world
120	5	5	it is important
121	5	5	it is the
122	5	4	marriage partner in
123	5	5	my opinion it
124	5	4	of their children
125	5	5	of them is
126	5	5	partner by yourself
127	5	5	right to make
128	5	5	some people believe
129	5	4	that is why
130	5	3	the age of
131	5	5	the choice of
132	5	4	their child s
133	5	5	their children and
134	5	5	their own choice
135	5	3	there are also
136	5	5	to choose the
137	5	4	to live with
138	5	4	when people choose
139	5	4	will be better
140	5	4	you will be

Appendix B – Three-word bundles of Anadolu University students

antconc result_anadolu_3_02.01.21 - Not Defteri

Dosya Düzen Biçim Görünüm Yardım

#Total No. of N-Gram Types: 202
#Total No. of N-Gram Tokens: 1405

1	36	26	arranged by parents
2	30	19	choose their own
3	24	23	in some countries
4	22	19	are arranged by
5	21	15	know each other
6	20	15	arranged marriages are
7	20	12	people should choose
8	19	13	don t know
9	19	18	marriages are arranged
10	18	16	people choose their
11	17	16	on the other
12	17	9	should choose their
13	16	15	the other hand
14	16	10	you don t
15	14	14	countries marriages are
16	14	14	some countries marriages
17	14	14	to sum up
18	13	6	i don t
19	12	10	their own partner
20	12	7	their own partners
21	12	3	with a partner
22	11	8	own marriage partner
23	11	9	they don t
24	11	7	to marry with
25	11	8	with a person
26	11	10	with someone you
27	10	10	are going to
28	10	10	by parents but
29	10	10	in my opinion
30	10	9	marriage is a
31	10	5	marry with a
32	10	10	parents but in
33	10	7	the person you
37	9	7	a lot of
38	9	9	but in other
39	9	8	choose their partner
40	9	7	in love with
41	9	8	it is not
42	9	6	it s not
43	9	7	most of the
44	9	8	the person who
45	9	7	their own marriage
46	8	6	an arranged marriage
47	8	6	be arranged by
48	8	8	countries people choose
49	8	8	first of all
50	8	8	in other countries
51	8	8	love each other
52	8	8	other countries people
53	8	6	to get married
54	8	4	we don t
55	8	3	with him her
56	8	5	you can t
57	7	7	as a result
58	7	6	choosing your own
59	7	5	don t want
60	7	4	going to marry
74	6	4	if there is
75	6	6	is an important
76	6	5	is the best
77	6	6	marriage is an
78	6	4	marry with the
79	6	5	on their own
80	6	4	one of the
81	6	6	they are going
82	6	4	to be happy
83	6	4	to marry someone
84	6	3	wife or husband
85	6	4	you didn t
86	6	4	you want to
87	5	3	a partner who
88	5	3	a person that
89	5	4	a result of
90	5	5	and it is
91	5	4	are a lot
92	5	4	arranged marriage is
93	5	4	because of the
94	5	4	by parents are
95	5	3	by the parents
96	5	4	can t know
97	5	5	don t love
98	5	5	for each other
99	5	5	for the rest
100	5	4	for their children
101	5	3	get married with

antconc result_anadolu_3_02.01.21 - Not Defteri

Dosya Düzen Biçim Görünüm Yardım

102	5	3	in arranged marriages
103	5	4	is one of
104	5	4	it is also
105	5	5	it should be
106	5	5	lead to a
107	5	4	marriage is arranged
108	5	3	of arranged marriage
109	5	5	of your life
110	5	4	should be arranged
111	5	5	that their parents
112	5	4	the people who
113	5	4	there are a
114	5	3	there are still
115	5	4	they do not
116	5	5	to each other
117	5	5	to have a
118	5	3	want to marry
119	5	5	we can t
120	5	4	whole life with
121	5	4	you do not
122	5	5	you marry with
123	5	4	you will marry
124	5	5	your own partner
125	5	5	your whole life
126	4	3	a person you
127	4	4	arrenged by parents
128	4	4	but i think
129	4	4	but it is
130	4	4	can say that
131	4	3	chance to know
138	4	4	he or she
139	4	3	how can you
140	4	3	husband or wife
141	4	4	i think it
142	4	4	i think people
143	4	4	i think that
144	4	3	in a relationship
145	4	4	in some cultures
146	4	4	in some parts
147	4	4	is not a
148	4	4	it comes to
149	4	4	it is important
150	4	3	it s a
151	4	4	life with a
152	4	4	life with someone
153	4	4	marriages are a
154	4	4	marriages are not
155	4	3	marriages arranged by
156	4	3	marriages that are
157	4	4	my opinion people
158	4	3	of arranged marriages
159	4	3	of our life
160	4	3	of the person
161	4	4	of the world
162	4	4	own partner is
163	4	4	person who is
164	4	4	rest of their
165	4	4	rest of your
166	4	3	should be a
167	4	3	should be the
176	4	3	that they can
177	4	4	the best for
178	4	3	the end of
179	4	3	the fact that
180	4	4	the most important
181	4	3	the one you
182	4	4	the person that
183	4	3	their children to
184	4	4	their parents arranged
185	4	3	they can t
186	4	3	they have to
187	4	4	think people should
188	4	3	think that the
189	4	4	this kind of
190	4	3	to be unhappy
191	4	4	to choose your
192	4	4	to spend your
193	4	4	type of marriage
194	4	4	up marriage is
195	4	3	we want to
196	4	4	when it comes
197	4	4	with someone who
198	4	3	you are married
199	4	4	you can not
200	4	3	you have to
201	4	3	you will be
202	4	4	your life with

Appendix C – Three-word bundles of Eskişehir Osmangazi University students

antconc result_esogu_3_02.01.21 - Not Defteri

Dosya Düzen Biçim Görünüm Yardım

#Total No. of N-Gram Types: 168
#Total No. of N-Gram Tokens: 1455

1	33	28	choose their own
2	26	22	arranged by parents
3	26	24	people choose their
4	25	21	in some countries
5	24	18	know each other
6	22	14	marriage is a
7	20	15	own marriage partner
8	20	16	their own marriage
9	19	16	on the other
10	19	16	the other hand
11	18	18	are arranged by
12	17	17	marriages are arranged
13	16	7	they do not
14	16	10	you don t
15	15	11	arranged marriage is
16	15	10	do not know
17	14	8	for their children
18	14	9	in arranged marriages
19	14	10	love each other
20	14	9	the most important
21	14	11	to each other
22	14	13	to know each
23	13	8	are going to
24	13	12	in my opinion
25	13	7	you do not
26	12	10	right to choose
27	11	11	countries marriages are
28	11	8	each other and
29	11	10	it is a
30	11	11	should choose their
31	11	9	the person who
32	11	10	their own partner
33	11	11	to sum up
34	11	10	want to marry
35	10	8	arranged marriages are
37	10	5	choose your partner
38	10	10	it can be
39	10	10	it is not
40	10	8	most of the
41	10	7	of your life
42	10	10	one of the
43	10	10	parents but in
44	10	9	people should choose
45	10	10	some countries marriages
46	10	10	the person you
47	10	7	the right to
48	10	10	they want to
49	10	9	your life with
50	9	5	a happy marriage
51	9	9	but in other
52	9	8	choose their partner
53	9	7	don t know
54	9	6	going to marry
55	9	7	there is no
56	9	5	to get married
57	9	7	you choose your
58	9	7	your own partner
59	8	6	an arranged marriage
60	8	8	arranged by the
61	8	7	do not love
62	8	8	fall in love
63	8	7	get to know
64	8	6	he or she
65	8	8	in other countries
66	8	8	is one of
67	8	6	it will be
68	8	4	know the person
69	8	6	not want to
70	8	6	of the people
71	8	4	the arranged marriages
72	8	7	the rest of
73	8	5	you are going
74	8	6	you get married
75	7	5	a lot of
76	7	7	choose the person
77	7	5	choosing your own
78	7	6	don t love
79	7	7	if they are
80	7	7	in love with
81	7	5	in order to
82	7	3	is better than
83	7	3	love marriage is
84	7	4	love your partner
85	7	7	marriage is the
86	7	4	marry with someone
87	7	7	other countries people
88	7	5	that s why
89	7	6	the right person
90	7	5	there will be
91	7	7	they are not
92	7	7	to choose their
93	7	7	to have a
94	7	5	with a person
95	7	7	with each other
96	7	6	you have to
97	6	5	according to their
98	6	6	at the end
99	6	4	be arranged by
100	6	5	choose their partners
101	6	6	countries people choose
102	6	6	for example in
103	6	5	get married and
104	6	4	him or her
105	6	6	i believe that
106	6	5	i think that
107	6	5	if you are
108	6	5	is the best

antconc result_esogu_3_02.01.21 - Not Defteri

Dosya Düzen Biçim Görünüm Yardım

109	6	5	is the most
110	6	6	it is the
111	6	6	it should be
112	6	4	marriages arranged by
113	6	6	of the most
114	6	4	on your own
115	6	5	partner for their
116	6	5	point of view
117	6	6	rest of your
118	6	6	some of them
119	6	5	there are many
120	6	6	there are two
121	6	6	there is a
122	6	6	to marry with
123	6	4	with your partner
124	6	5	your partner you
125	5	5	a life with
126	5	5	a long time
127	5	5	a person who
128	5	3	a person whom
129	5	5	are forced to
130	5	4	arranged marriage or
131	5	4	as a result
132	5	5	by parents but
133	5	5	by your parents
134	5	5	cases people choose
135	5	5	chosen by your
136	5	5	do not want
137	5	5	each other before
138	5	5	each other better
147	5	5	marriage can be
148	5	5	marriage partner in
149	5	5	marriage should be
150	5	4	marry someone you
151	5	5	of their lives
152	5	4	own partner is
153	5	5	people do not
154	5	4	people who will
155	5	4	pros and cons
156	5	5	rest of their
157	5	5	someone you love
158	5	5	that marriage is
159	5	3	the best for
160	5	4	the rate of
161	5	4	they are going
162	5	5	to choose the
163	5	4	to live with
164	5	4	to marry to
165	5	3	when you are
166	5	3	with that person
167	5	4	you can not
168	5	4	you marry someone

Appendix D – Three-word bundles of native students

antconc result_LOCNESS_3_07.01.21 - Not Defteri

Dosya Düzen Biçim Görünüm Yardım

#Total No. of N-Gram Types: 222
#Total No. of N-Gram Tokens: 8047

1	163	11	the fact that
2	130	12	in order to
3	123	13	one of the
4	118	5	the united states
5	106	12	that it is
6	95	9	be able to
7	94	13	there is no
8	90	3	icle br sur
9	85	9	the right to
10	83	10	it is not
11	82	12	due to the
12	82	8	the end of
13	80	9	because of the
14	77	9	the idea of
15	77	13	there is a
16	76	10	as well as
17	70	9	end of the
18	70	12	it is a
19	69	11	the use of
20	69	12	this is a
21	68	10	of the world
22	67	6	in the united
23	66	8	should not be
24	65	8	the number of
25	64	13	in the world
26	64	10	to be a
27	62	11	it is the
28	61	11	a lot of
29	61	11	that they are
30	59	11	part of the
31	58	8	on the other
32	57	12	it would be
37	49	3	for the best
38	47	11	there are many
39	47	6	they do not
40	47	7	to have a
41	46	8	is not the
42	45	11	as a result
43	45	9	out of the
44	44	3	invention of the
45	44	5	the invention of
46	43	11	such as the
47	43	8	the rest of
48	43	9	to be the
49	42	4	that he is
50	41	5	the question of
51	40	6	all of the
52	40	11	this is the
53	39	9	is not a
54	39	8	is one of
55	39	12	it has been
56	39	7	the case of
57	39	5	the th century
58	39	9	this is not
59	38	9	aware of the
60	38	5	the amount of
61	38	4	the beginning of
62	38	9	there would be
63	37	9	but it is
64	37	9	of the most
65	37	10	the majority of
66	37	9	would be a
67	36	7	because it is
68	36	9	in the past
69	35	11	more and more
70	35	10	some of the
71	35	10	there has been
72	34	10	can be seen
73	34	7	in the case
74	34	6	in the s
75	34	8	is that the
76	34	9	most of the
77	34	8	would not be
78	33	5	a part of
79	33	8	a result of
80	33	8	as a whole
81	33	11	as it is
82	33	8	i feel that
83	33	8	is a very
84	33	8	it can be
85	33	8	many of the
86	33	9	that there is
87	32	9	i believe that
88	32	8	i think that
89	32	9	the introduction of
90	32	3	the prime minister
91	31	3	a loss of
92	31	6	as long as
93	31	6	because they are
94	31	7	would have to
95	30	5	according to the
96	30	6	an example of
97	30	3	he or she
98	30	7	in the future
99	30	3	of the th
100	30	10	the only way
101	30	6	the role of
102	30	6	the u k
103	30	7	to say that
104	30	7	whether or not
105	29	8	and it is
106	29	6	fact that the
107	29	5	he does not
108	29	9	in my opinion

antconc result_LOCNESS_3_07.01.21 - Not Defteri

Dosya Düzen Biçim Görünüm Yardım

109	29	5	of the people
110	29	5	that this is
111	29	7	they don t
112	29	8	to make the
113	28	6	be allowed to
114	28	7	it is possible
115	28	8	the ability to
116	28	7	the effects of
117	28	10	there will be
118	28	4	this type of
119	28	7	to make a
120	27	10	are able to
121	27	8	believe that the
122	27	3	his or her
123	27	5	not want to
124	27	3	the european community
125	26	10	a great deal
126	26	9	at the same
127	26	11	it is also
128	26	9	people who are
129	26	9	seems to be
130	26	6	the death of
131	26	6	the development of
132	26	7	the issue of
133	26	7	the world and
134	26	5	they want to
135	26	6	up to the
136	26	5	way of life
137	25	4	a sense of
147	25	5	the loss of
148	25	9	the same time
149	25	4	the u s
150	25	8	to the fact
151	25	9	will not be
152	24	9	away from the
153	24	7	be used to
154	24	3	br sur the
155	24	8	in a way
156	24	5	is for the
157	24	6	it does not
158	24	8	would like to
159	23	10	for example the
160	23	7	is in the
161	23	9	may not be
162	23	3	the futility of
163	23	4	the most important
164	23	8	the problem of
165	23	6	to a certain
166	23	3	to have children
167	22	8	an increase in
168	22	8	be seen as
169	22	6	do not have
170	22	7	i don t
171	22	7	in order for
172	22	5	in the uk
173	22	6	in which the
174	22	10	it could be
175	22	7	it is an
176	22	3	of the computer
177	22	6	only way to
178	22	5	that of the
179	22	7	the idea that
180	22	6	the lives of
181	22	7	the power of
182	22	4	they are not
183	22	7	to do with
184	22	7	will have to
185	21	5	a matter of
186	21	4	amount of money
187	21	8	being able to
188	21	5	example of this
189	21	6	have the right
190	21	7	i do not
191	21	10	in the long
192	21	7	in the u
193	21	7	is not only
194	21	4	it is important
195	21	8	it is only
196	21	9	it will be
197	21	7	majority of the
198	21	7	rest of the
199	21	7	seem to be
200	21	6	the risk of
201	21	6	there have been
202	21	7	there should be
203	21	7	to go to
204	21	5	when it comes
205	20	10	a number of
206	20	5	are going to
207	20	10	at the moment
208	20	9	great deal of
221	20	5	will continue to
222	20	4	would not have
212	20	3	needs to be
213	20	4	seen to be
214	20	6	that they can
215	20	5	the age of
216	20	7	the form of
217	20	5	the united kingdom
218	20	4	to deal with
219	20	6	to do this
220	20	7	to take the

Appendix F – Bundles of Turkish students compared to bundles of native students

N	Key word	Freq.	%	Texts	RC. Freq.	RC. %	Keyness	P	Lemmas	Set
1	CHOOSE THEIR OWN	63	0,18	47	0		295,65	0,00		
2	ARRANGED BY PARENTS	62	0,18	48	0		290,96	0,00		
3	IN SOME COUNTRIES	49	0,14	44	0		229,94	0,00		
4	KNOW EACH OTHER	44	0,13	32	0		206,47	0,00		
5	PEOPLE CHOOSE THEIR	44	0,13	40	0		206,47	0,00		
6	ARE ARRANGED BY	40	0,12	37	0		187,69	0,00		
7	MARRIAGES ARE ARRANGED	36	0,10	35	0		168,92	0,00		
8	MARRIAGES ARE ARRANGED BY	33	0,10	32	0		154,84	0,00		
9	MARRIAGE IS A	32	0,09	23	0		150,15	0,00		
10	PEOPLE CHOOSE THEIR OWN	32	0,09	29	0		150,15	0,00		
11	OWN MARRIAGE PARTNER	31	0,09	23	0		145,45	0,00		
12	ARRANGED MARRIAGES ARE	30	0,09	23	0		140,76	0,00		
13	ARE ARRANGED BY PARENTS	30	0,09	28	0		140,76	0,00		
14	PEOPLE SHOULD CHOOSE	30	0,09	21	0		140,76	0,00		
15	THEIR OWN MARRIAGE	29	0,08	23	0		136,07	0,00		
16	SHOULD CHOOSE THEIR	28	0,08	20	0		131,38	0,00		
17	TO SUM UP	25	0,07	25	0		117,30	0,00		
18	CHOOSE THEIR OWN MARRIAGE	25	0,07	22	0		117,30	0,00		
19	COUNTRIES MARRIAGES ARE	25	0,07	25	0		117,30	0,00		
20	TO KNOW EACH OTHER	24	0,07	21	0		112,61	0,00		
21	OME COUNTRIES MARRIAGES ARE	24	0,07	24	0		112,61	0,00		
22	SOME COUNTRIES MARRIAGES	24	0,07	24	0		112,61	0,00		
23	THEIR OWN MARRIAGE PARTNER	24	0,07	18	0		112,61	0,00		
24	TO KNOW EACH	24	0,07	21	0		112,61	0,00		
25	IN SOME COUNTRIES MARRIAGES	24	0,07	24	0		112,61	0,00		
26	PEOPLE SHOULD CHOOSE THEIR	23	0,07	16	0		107,91	0,00		
27	LOVE EACH OTHER	22	0,06	18	0		103,22	0,00		
28	RIES MARRIAGES ARE ARRANGED	22	0,06	22	0		103,22	0,00		
29	THEIR OWN PARTNER	22	0,06	19	0		103,22	0,00		
30	DO NOT KNOW	21	0,06	15	0		98,53	0,00		
31	ARRANGED MARRIAGE IS	20	0,06	15	0		93,84	0,00		
32	PARENTS BUT IN	20	0,06	20	0		93,84	0,00		
33	FOR THEIR CHILDREN	19	0,06	12	0		89,14	0,00		
34	YOU DONÂ T	19	0,06	12	0		89,14	0,00		
35	CHOOSE THEIR OWN PARTNER	19	0,06	16	0		89,14	0,00		
36	IN ARRANGED MARRIAGES	19	0,06	12	0		89,14	0,00		
37	SHOULD CHOOSE THEIR OWN	18	0,05	13	0		84,45	0,00		
38	THE PERSON WHO	18	0,05	16	0		84,45	0,00		
39	TO EACH OTHER	18	0,05	15	0		84,45	0,00		
40	BUT IN OTHER	18	0,05	18	0		84,45	0,00		
41	WITH A PERSON	18	0,05	13	0		84,45	0,00		

Appendix F – Bundles of Kazakh students compared to bundles of native students

N	Key word	Freq.	%	Texts	RC. Freq.	RC. %	Keyness	P	Lemmas	Set
1	OWN MARRIAGE PARTNER	41	0,22	29	0		238,53	0,00		
2	CHOOSE THEIR OWN	39	0,21	29	0		226,89	0,00		
3	PEOPLE CHOOSE THEIR	35	0,19	29	0		203,62	0,00		
4	THEIR OWN MARRIAGE	33	0,18	28	0		191,98	0,00		
5	THEIR OWN MARRIAGE PARTNER	33	0,18	28	0		191,98	0,00		
6	CHOOSE THEIR OWN MARRIAGE	32	0,17	27	0		186,16	0,00		
7	IN SOME COUNTRIES	31	0,17	26	0		180,34	0,00		
8	BY THE PARENTS	30	0,16	26	0		174,52	0,00		
9	PEOPLE CHOOSE THEIR OWN	29	0,16	25	0		168,70	0,00		
10	ARRANGED BY THE	26	0,14	22	0		151,25	0,00		
11	ARRANGED BY THE PARENTS	25	0,13	22	0		145,43	0,00		
12	BUT IN OTHER	24	0,13	24	0		139,61	0,00		
13	FOR THEIR CHILDREN	22	0,12	18	0		127,97	0,00		
14	ARE ARRANGED BY	21	0,11	18	0		122,16	0,00		
15	PARENTS BUT IN	21	0,11	21	0		122,16	0,00		
16	I WANT TO	20	0,11	19	0		116,34	0,00		
17	MARRIAGES ARE ARRANGED	19	0,10	18	0		110,52	0,00		
18	ARE ARRANGED BY THE	19	0,10	17	0		110,52	0,00		
19	MARRIAGES ARE ARRANGED BY	19	0,10	18	0		110,52	0,00		
20	PARENTS BUT IN OTHER	19	0,10	19	0		110,52	0,00		
21	IN OTHER CASES	18	0,10	18	0		104,70	0,00		
22	IN THIS ESSAY	18	0,10	17	0		104,70	0,00		
23	THE PARENTS BUT	17	0,09	17	0		98,88	0,00		
24	THE PARENTS BUT IN	17	0,09	17	0		98,88	0,00		
25	CASES PEOPLE CHOOSE THEIR	17	0,09	17	0		98,88	0,00		
26	CASES PEOPLE CHOOSE	17	0,09	17	0		98,88	0,00		
27	BY THE PARENTS BUT	17	0,09	17	0		98,88	0,00		
28	OTHER CASES PEOPLE CHOOSE	16	0,09	16	0		93,07	0,00		
29	BUT IN OTHER CASES	16	0,09	16	0		93,07	0,00		
30	IN OTHER CASES PEOPLE	16	0,09	16	0		93,07	0,00		
31	OTHER CASES PEOPLE	16	0,09	16	0		93,07	0,00		
32	TO SUM UP	15	0,08	15	0		87,25	0,00		
33	THIS ESSAY I	14	0,07	14	0		81,43	0,00		
34	THIS KIND OF	14	0,07	13	0		81,43	0,00		
35	IN THIS ESSAY I	14	0,07	14	0		81,43	0,00		
36	COUNTRIES MARRIAGES ARE	13	0,07	13	0		75,61	0,00		
37	SOME COUNTRIES MARRIAGES	13	0,07	13	0		75,61	0,00		
38	OME COUNTRIES MARRIAGES ARE	13	0,07	13	0		75,61	0,00		
39	PARTNER FOR THEIR	13	0,07	11	0		75,61	0,00		
40	IN SOME COUNTRIES MARRIAGES	13	0,07	13	0		75,61	0,00		
41	PURPOSE OF THIS ESSAY	12	0,06	12	0		69,80	0,00		

Appendix G – Bundles of native students compared to bundles of Turkish and Kazakh students

N	Key word	Freq.	%	Texts	RC. Freq.	RC. %	Keyness	P	Lemmas	Set
1	THE UNITED STATES	118	0,04	5	0		35,78	0,00		
2	IS THE MOST	9		5	15	0,03	-29,82	0,00		
3	THEY WANT TO	26		5	23	0,04	-30,37	0,00		
4	IN CONCLUSION I	7		4	15	0,03	-33,45	0,00		
5	POINT OF VIEW	8		6	21	0,04	-50,65	0,00		
6	FIRST OF ALL	16		5	26	0,05	-51,03	0,00		
7	ARE GOING TO	20		5	29	0,05	-53,57	0,00		
8	EACH OTHER AND	9		4	24	0,05	-58,21	0,00		
9	THE MOST IMPORTANT	22		4	32	0,06	-59,22	0,00		
10	ON THE OTHER	58	0,02	8	56	0,11	-79,30	0,00		
11	ON THE OTHER HAND	50	0,02	6	54	0,10	-83,01	0,00		
12	THE OTHER HAND	50	0,02	6	55	0,10	-85,63	0,00		
13	IN MY OPINION	29		9	46	0,09	-89,19	0,00		

Appendix H – Sample essay N1

Even nowadays there are societies in which the main role of establishing the marriage is acted by parents permission, however mainly in progressive countries people have more freedom for they are private lives. Conservative members of world community state that young generation isn't able to choose their life partner wisely and against them standing the owners of liberal views who assert that every human should have his own choice, it's his right. People are equally able to select the path of life and companion, so the pressure on them is unacceptable. But everything depends on cases and people relationships.

We have the practice of parents invasion into their children private lives in our country too. Mostly parents just want to "help" their son (doughter) to find his (her) place in life, because being married and have 3 kids is a standart, but we can understand them - they want grandchildren. Nevertheless, as for me such actions are ancient. All people are different and each of them picking own way in this world, so forcing them to do something with their lives without will is cruel.

On other hand is respect to humans and lifestyle that they have chosen. Most of humanity believes that it's really important not to judje people for made decisions, if they want to spend life for career or family (doesn't matter) we haven't got any rights to codemn them, although in some cases an individual needs help in searching a partner, so here the first opinion is fair. I've no any objection on second statement. Because the right for self management is one of the prior rights.

To sum up both opinions have own reasons, advantages and disadvantages. Even so second view is more fair on world of tolerance background.

Appendix I – Sample essay N2

Marriage. The most important event in the life of every person. After this your lifestyle may change forever. Some of us understand this, but others can argue with this statement or just tend to think that it is not so relevant. So the first part of people usually prefer to choose their soulmates by themselves, but there are also those, who allows their parents to choose their spouses instead of them or just obey to parent's decision.

Choosing your soulmate by yourself is brave and powerful act. You should do it, because YOU will live with this person, not your parents. You should respect him, if you want your love to be long, but for it first of all you should to know well this person. And, finally, if something will be wrong with your relationship, you won't judge others because of it. You will always know that it is your choice. But on the other hand, you could make the biggest mistake in your life and it could be better if someone will help you in this period of your life.

In the case when your marriage will be arranged by your parents there are also proms and cons. One of the advantages of it is that your parents know you very well and they just can't make a mistake, however there are also the risk that they can. Another positive side is that it would be easier to you to find your spouse instead of looking for him for a long time, if, of course, your parents will succeed to find the best variant of your life partner.

My Mom and Dad found each other by themselves and now they are happy together. My uncle and his wife married because of the decision of their parents and now they are happy together too. So, getting all into account, I can state that everything depends on the fortune.

And I am still tend to believe that the true love between two people are determined and their combined purpose is just to find each other.

Appendix J

Ethics Committee Approval

Evrak Kayıt Tarihi: 19.02.2019	Protokol No: 14836	Tarih: 28.03.2019

ANADOLU ÜNİVERSİTESİ
SOSYAL VE BEŞERÎ BİLİMLER BİLİMSEL ARAŞTIRMA VE YAYIN ETİĞİ KURULU
KARAR BELGESİ

ÇALIŞMANIN TÜRÜ:	Yüksek Lisans Tez Çalışması
KONU:	Eğitim Bilimleri
BAŞLIK:	Öğrenci Yazımında Sözcük Öbekleri: Türk ve Kazak İngilizce Öğretmenliği Bölümü Öğrencilerinin Tartışmacı Metinlerinin Analizi Lexical Bundles in Learner Writing: An Analysis of English Argumentative Essays of Turkish and Kazakhstani ELT students
~~PROJE~~/TEZ YÜRÜTÜCÜSÜ:	Prof. Dr. Gül DURMUŞOĞLU KÖSE
TEZ YAZARI:	Madina YÖRÜK
ALT KOMİSYON GÖRÜŞÜ:	-
KARAR:	Olumlu

Prof.Dr. Coşkun BAYRAK *(Başkan-Eğitim Fak.)*	
Prof.Dr. T. Volkan YÜZER *(Başkan Yardımcısı-Açıköğretim Fak.)*	**Prof.Dr. Esra CEYHAN** *(Eğitim Fak.)*
Prof.Dr. Münevver ÇAKI *(Güzel Sanatlar Fak.)*	**Prof.Dr. M. Erkan ÜYÜMEZ** *(İkt. ve İdari Bil. Fak.)*
Prof.Dr. Handan DEVECİ *(Eğitim Fak.)*	**Prof.Dr. Emel ŞIKLAR** *(İkt. ve İdari Bil. Fak.)*