



## A bibliometric analysis of lexical bundles usage in native and non-native academic writing

Fatma Demiray Akbulut<sup>a 1</sup> 

<sup>a</sup> Bolu Abant İzzet Baysal University, Bolu, Turkey

### APA Citation:

Akbulut, D., F. (2020). A bibliometric analysis of lexical bundles usage in native and non-native academic writing. *Journal of Language and Linguistic Studies*, 16(3), 1146-1166. Doi: 10.17263/jlls.803583

Submission Date:26/12/2019

Acceptance Date:01/07/2020

### Abstract

Over the last few decades, it has become obligatory to publish academic research to survive in academic world. In academic writing skill, lexical bundles (LBs) usage is marked as a master ability to build well-organized paragraphs or essays and are essential instruments to sound more natural and be fluent in English. The current study aimed to search the usage of LBs by non-native and native English academic writers and comparing the profiles of LBs in terms of the function, structure, and frequency. A corpus of 257 articles written by native and non-native academic writers was organized and multi-word LBs were identified using AntConc 3.2.4w. Written texts are collected from Journal of Language and Linguistic Studies (JLLS) by non-native speakers (NNSs) and Linguistics and Education Journal (LINGED) by native speakers (NSs). The findings suggest that non-native writers generally use more lexical bundles but fall into more repetitions. Besides, there are some substantial differences in the frequencies of structural and functional usage of LBs. In structural category, NNSs used Noun Phrase (NP) and Prepositional Phrase (PP)-based LBs and Conjunction at a lower rate, and Verb Phrase (VP) and Clause-based LBs at a higher rate than NSs. In terms of functional category, research-oriented LBs are more frequently used by NSs while text and stance-oriented LBs are more frequently used by NNSs. As a conclusion, it is stated that, the lexico-grammatical difference between two languages and the NNSs' tendency to write result-oriented article revealed distinct features of LB usage.

© 2020 JLLS and the Authors - Published by JLLS.

*Keywords:* lexical bundles; academic writing; multi-word units; lexical teaching; English language

## 1. Introduction

The formulaic language is an umbrella term used to refer to multi-word units in a foreign language, as well as elements that form the basis of a language. These structures, which take their place in the literature as multi-word expressions, are also one of the most important parts of fluent linguistic production. Basically, we can say that these lexical structures are linguistic regulators that follow each other frequently, provide integrity within the text and ensure the distinguishability of contextual

<sup>1</sup> Corresponding author. Tel.: +90-506-471-7261  
E-mail address: demiray\_f@ibu.edu.tr

structures. In other words, as Hyland (2008a) mentions, in order to gain control of a new language, users need to gain sensitivity in using certain lexical structures more than others.

Formulaic language structures which can be seen as a part of phraseology include multi-word units, idiomatic structures, repetitive sequences, lexical bundles or collocations. With the advancement of computer-mediated research methodology, studies in different disciplines and genres are developing rapidly within the context of phraseology. In addition, studies in this field are becoming part of both linguistics and applied linguistics. Corpus data has been used in the literature especially in the field of linguistics, psycholinguistics and foreign language teaching in order to carry out studies related to formulaic language structures. For example, the structure of “as can be seen” is frequently encountered in academic writing, while “as you can see” or “as can be observed” structures are rarely encountered, which represents a psychological relationship between words and a real relationship between users' communicative experiences (Hyland, 2008a: 5). In recent years, especially in countries such as the United States where English is the mother tongue or dominated, the majority of studies focuses on the writing of university students (Staples & Reppen, 2016). In addition to the studies focusing on the second language academic writing, the studies focusing on LBs or repetitive multi-word structures used in written language (Biber, Conrad, & Cortes, 2004; Cortes, 2004), comparison of academic articles written in mother tongue and second language (Adel & Erman, 2012; Bychkovska & Lee, 2017; Chen & Baker, 2010; Huang, 2015) comparative and bibliometric analysis of these structures (Chen & Baker, 2016; Lu & Deng, 2019; Shin, 2019) are also included. When we look at the study related to phraseology in the past by Ang, Tan and He (2017), it can be understood that it is necessary to know these structures for natural and fluent use of language. As a paramount component of a contentful text, LBs are very common in language, but it might be a complicated issue to find the appropriate LB for non-native writers. The widespread use of idiomatic sequences in discourse depends on the storage of phrasologic structures based on creating and understanding meaning in the lexicon of language user. Within the academic discourse, the expertise gained in the relevant phrasal structures is important because it enables academic writers to reach the relevant academic community.

## *1.1. Review of Literature*

### *1.1.1. Defining lexical bundles*

The lexical bundle is one of the features of an academic prose that can be defined as recurrent sequences of vocabulary items (Biber & Barbieri, 2007) or a combination of repeated three or more words in a specific corpus and in the same form and sequence (Coxhead & Byrd, 2007: 137). Biber (2006) notes the importance of LBs reporting high-frequency multi-word sets in the last decade. The diversity of the discursive functions and meanings of the LBs seems to be beneficial for academic writers, and these benefits can be expressed as follows: The first benefit, LBs makes it easier for advanced students to write because students can make sentences with ready-made parts easier than dealing with individual words. A second benefit is that these LBs are often used in academic writing, helping students to practice writing fluently. The last and perhaps the most important benefit is that these structures allow students to see the different elements between syntax and lexis (Biber, 2006; Coxhead & Byrd, 2007).

In the field of language learning, the importance of LBs is increasing in pedagogical terms. Knowing and distinguishing the word sequence can be considered as important to be able to perceive word sequences as a whole. In addition to the abovementioned studies, there is also evidence in the literature that holistic knowledge of word sequences improves processing in language learners (Siyanova-Chanturia, Conklin, & Van Heuven, 2011). In addition, studies in which LBs have been examined according to different types of discourse (eg Biber, Conrad, & Cortes, 2004; Grabowski, 2015), studies on phrasal or formulaic language which were used in different English teaching materials (Meunier & Gouverneur, 2007) show the importance of LBs' usage.

## *1.2. Lexical Bundles in Academic Writing*

In their study, Chen and Baker (2010) described an approach that identified commonly used word combinations in academic writing, such as lexical bundles. One of the corpora used in their study consisted of published academic texts and the other one consisted of academic writings of students (L1 and L2, respectively). While LBs were rarely used in second language students' academic writing, it was seen that they were used frequently in academic writing of L1 users. In addition, L2 students used some structures such as "all over the world" quite often, while some high-frequency words were used less frequently in academic writings of students than in published academic texts. Juknevičienė (2009) compared the language produced by students in three different proficiency levels on the basis of LBs. As a result of the comparison of corpora from written texts of the students, it was shown that low-level students repeated their LBs structures more. Because of the low levels, students with more limited LBs structures fall into more repetitions. Concerning LBs usage, Bychkovska and Lee (2017) examined the essays written by two groups of university students (L1 English and L1 Chinese). In a total of 206 article studies (101, L1 English students' essays and 105 L1 Chinese students' essays), the findings showed that L2 students used more LBs than L1 students, but the usage of structural and functional bundle types was different from each other. L1 writers generally used noun and prepositional phrases, while L2 writers used more verb phrases.

In their study, Ang and Tan (2018) conducted a corpus-oriented study by examining the use of LBs in academic writing in the context of academic purpose-based English teaching. One of them consisted of journal articles written in the field of International Business Management (IBM), while the other consisted of LBs in the Academic Formulas List (AFL). According to the results, three-word LBs were widely used in IBM corpus. Therefore, it was considered beneficial to demonstrate a LBs-oriented approach and teaching method within the EAP environment. Similarly, Ghafar Samar, Shokrpour and Nasiril (2018) investigated the applicability of LBs teaching in the writing performance of medical students in EFL context experimentally (pre-test, treatment, post-test). According to the results of the study, the use of LBs changed significantly after the treatment and medical students started to use LBs more naturally and native-like. Another study examining the usage of LBs in written texts belongs to Ruan (2017). The researcher investigated Chinese university students learning L2 English. In the study, frequency-based bundles were examined using academic text corpus that students wrote every year from 1st to 4th year. According to the results, students used more and different types of LBs as time passed, and also the structural forms of preferred LBs changed. Biber and Barbieri (2007) found a result that was contrary to some previous studies and concluded that more LBs were used in spoken language than written language. In their study, a wide range of written and spoken university registers were examined and the results showed that more LBs were used in the written language instead of spoken language. Cortes (2013) examined the LBs used in the introductory chapters of research articles and discovered that "a group of LBs were exclusively linked to one move or step in a move while a second group occurred across several moves and steps" (p.33). Shin (2019) also examined the LBs used in university students' articles. These students are composed of native and non-native English speakers. The study examined the structures and functions of LBs used in corpora. The researcher concluded that both groups similarly used LBs in the context of VP-based bundles, stance-expression bundles, idiomatic PP bundles, and informal quantifying bundles. Lu and Deng (2019) conducted a study examining the four-word LBs used in the abstract parts of the theses written in English by L1 Chinese students at Tsinghua University and L1 English students at the Massachusetts Institute of Technology (MIT Corpus). LBs were categorized as structural and functional groups. According to the results, missing knowledge of some features in English lexico-grammatical system was observed in Chinese students' writing. Also, they were not aware of the characteristic of bundles in hard science disciplines as much as L1 English students. In addition to these studies, in the literature there are a number of studies which examined the

effect of L1 on the use of LBs (Paquot, 2017), analysed the use of LBs in university books and textbooks (Biber, Conrad & Cortes, 2004), showed the usage and importance of LBs in different disciplines and literature (Hyland, 2008; Wright, 2019), identified the effect of LBs on reading skills (Allan, 2016; Tremblay, Derwing, Libben, Westbury, 2011) and demonstrated the LB use of academic lecturers (Nesi & Basturkmen, 2006).

### *1.3. Lexical Bundles in Turkish EFL Context*

There is a limited number of studies that have conducted corpus oriented lexical bundles in EFL context. At the beginning of these studies, Bal (2010) investigated the use of LBs in academic studies written in English by Turkish authors. In the study which examined the articles from six different academic disciplines, Bal (2010) concluded that LBs had structural correlation as well as strong functional properties. Similarly, Karabacak and Qin (2013) examined the LBs in English written argumentative essays which they gathered from L1 Turkish, L1 Chinese and L1 English writers. The results showed that even individuals with advanced English might not be able to fully comprehend these LBs and as a conclusion, they expressed that an explicit learning based on LBs was necessary.

Concerning LBs studies in Turkish EFL context, Öztürk and Köse (2016) examined the use of four-word LBs in the discipline of foreign language learning. They searched L1 Turkish and English postgraduate students' corpora and their usage of LBs. They investigated frequency, function and structure properties of these bundles. According to the results, Turkish postgraduate students used LBs more frequently than L1 English writers in their articles. However, when we looked at the number of LBs, it was said that Turkish students used the same LBs quite frequently. In terms of structural and functional categories, there was no significant difference between the groups. Güngör and Uysal (2016) also made a structural and functional analysis of LBs usage in academic writing in English by NNSs. Their findings showed that Turkish scholars overused the LBs in research articles. Considering the findings acquired within the Turkish context, it is obvious that the overuse of LBs is certainly remarkable. Finally, Uçar (2017) tried to examine the most commonly used three-word LBs in published English articles by L1 Turkish and L1 English writers and looked for differences in frequency, structure or function. 15 articles written by NSs and 15 articles written by NNSs of English were examined. According to the results, it was concluded that the use of LBs by Turkish writers was less varied and used less frequently than NSs.

### *1.4. Present Study*

Because of the limited number of articles on comparative analyses written on LBs in Turkish literature, the present study identifies published journal articles written by L1 Turkish and L1 English academicians or writers. For this purpose, the corpus was tried to be kept wide and four-word LBs were examined in terms of frequency, function and structure in a total of 264 articles. The LBs list was taken from Lu and Deng's (2019) study and was modified according to the pmw results. In other words, the LBs which were appeared less than 10 times per million words (Biber et al., 1999:137) were omitted from the list and after modification, the last version of LBs list was used (see Appendix A.). The reason for choosing four-word LBs is to ensure that they give clearer range of functional and structural features than three-word LBs (Hyland, 2008b, p.8). Three-word LBs have also been identified as the shorter versions of four-word LBs (Cortes, 2004). In this respect, three-word LBs have not been added to the list. The main aim of this study is to examine the frequency, function and structure of four-word LBs used in published academic articles in the years of 2017, 2018 and 2019 by L1 Turkish and L1 English authors. Based on this aim, the main research questions of the study can be listed as follows:

1. What are the types of most frequent four-word LBs used in articles in the discipline of language and linguistics by native and non-native academic writers?
2. Does the structure of LBs used in both groups differ from each other?
3. Does the function of LBs used in both groups differ from each other?

## 2. Method

### 2.1. Corpora

This study used a corpus that contains particular articles mostly about language and linguistics published between the years of 2017 and 2019. The first part of the corpora used in the current study includes the articles published in “*Journal of Language and Linguistic Studies (JLLS)*” which ranks the research articles on the website of [dergipark.gov.tr](http://dergipark.gov.tr) in Turkey. These articles were published in English by L1 Turkish writers. The second part of the corpora includes articles published in “*ScienceDirect Linguistics and Education (LINGED)*” journal by NSs of English. The main reason for the selection of both journals is that the published articles are research articles and that the topics are similar to each other in terms of focusing on linguistics and language education. Since both journals are Open-Access, the articles are easily accessible from the journals’ websites. While examining the studies in JLLS journal, articles published by Turkish authors were taken, and articles of foreign authors were eliminated from the list. Thus, a total of 157 articles and 1.007.812 words were analysed. Similarly, the writers whose native language cannot be English were excluded from the studies in the LINGED journal, and the articles of the NSs were examined by looking at the writers’ names and countries of the author and obtaining biographical information about the authors. In total, 107 articles and 1.006.648 words were analysed from LINGED journal. Instead of the number of articles, the number of words was taken into consideration because the average number of words in articles in LINGED was 9,408, while the average number of articles in JLLS was 6,419. Instead of reducing the word counts of the articles in the LINGED journal, the word numbers were tried to be synchronized and for this purpose, some articles from the LINGED journal were excluded from the study. The contents of the articles have similar characteristics and generally consists of Abstract, Introduction, Literature Review, Method (Participants, Instruments, Data Collection and Analysis), Results or Findings, Discussion and Conclusion. Table 1 represents the corpus of the study, the numbers of texts and total number of words they included.

**Table 1.** Corpora Description

	JLLS Corpus	LINGED Corpus
Number of Articles	157	107
Range of Length of Articles	2,279-12,960	2,779-12,656
Mean Length of Articles	6,419	9,408
Total words number	1.007.812	1.006.648

In this regard, 157 research articles written in English by non-native English scholars and 107 articles by native scholars (with a total of 2.014.460 words) were collected considering namely topic and text length to compare the frequency and structure of LBs. In total, LBs extracted from 264 articles were examined via AntConc (Anthony, 2019). The LBs frequencies for both corpora were tested for statistical significance using Paul Rayson’s log-likelihood calculator from the website of <http://ucrel.lancs.ac.uk/llwizard.html>. This analysis type is used “as a quick way in to find the differences between the corpora and is shown to have applications in the study of social differentiation

in the use of English vocabulary, profiling of learner English and document analysis in the software engineering process” (Rayson & Garside, 2000, p.1). In this analysis, two corpora were chosen as in the present study, a frequency list was produced (i.e. word level in the current study) and loglikelihood statistic was calculated for each word in this list.

### 3. Results

#### 3.1. Lexical Bundles Usage in Native and Non-native Corpora

To answer the first research question, LBs in academic essays by native and non-native writers were identified. In Table 2, all types of LBs from both groups (NNSs: 140 types, NSs: 132 types) have been demonstrated. In total, 131 LBs were used by both groups, it means that most of the LBs are not unique to only one group. In addition, LBs used by both groups have been analysed according to main structural types such as NP-based, PP-based, VP-based, Clause-based and Conjunctions and functional types such as Research, Text and Stance. As can be seen in the table, there is not much variation between two groups in terms of types of LBs used by NNSs and NSs. Specifically, LBs based on NP (NNSs and NSs: 45 types) and PP (NNSs and NSs: 44 types) were used equally by both groups in terms of types. When VP-based LBs were examined, it was seen that NNSs (22 types) used more LBs than NSs (18 types). Similarly, Clause-based LBs were used more extensively by NNSs (NNSs: 23 types, NSs: 19 types) and they used more different types of LBs. In terms of functional analysis (see Table 6), the general results show that NNSs used more types of functional LBs than NSs in a similar way to structural types (NNSs: 140 types, NSs: 133 types). The biggest difference between two groups can be seen in the type of *Resultative* LBs which is the subcategory of text-oriented LBs.

#### 3.2. Structural Analysis of Lexical Bundles

To answer the second research question, structural analysis of LBs in terms of NP-based, PP-based, VP-based, Clause-based and Conjunctions have been analysed in detail. Different structural LBs’ usage by native and non-native writers have been summarized in Table 2 and the proportion of them have been demonstrated in Table 3 and Figure 1. When Table 3 and Figure 1 are examined, it is seen that NNSs used NP and PP-based LBs and Conjunction at a lower rate, and VP-based and Clause-based LBs at a higher rate than NSs. Both groups used NP-based LBs in a close proportion level to each other, but NSs preferred a higher proportion of PP-based LBs. Starting from this point, as mentioned in the literature, while phrasal bundles are indicators of academic field, clausal bundles are more commonly used structures in colloquial language (Chen & Baker, 2016). NP and PP-based LBs are evaluated as phrasal bundles, while VP, Clause-based LBs and conjunctions are clausal LBs (Shin, 2019: 5). The study of Shin (2019) shows that in both NNSs and NSs corpus, phrasal bundles such as NP or PP-based are used less frequently. However, VP-based LBs, ie clausal bundles, were used more by both groups. However, the current study shows different results from Shin’s (2019) study that, both NNSs and NSs used phrasal bundles (NP and PP-based) more than clausal bundles (VP, Clause-based and Conjunctions). It can be said that as indicators of academic field, phrasal bundles are more preferred by academic writers in written context.

**Table 2.** The Categories and Subcategories of Structural Lexical Bundles

Category	Subcategory	Types		Tokens		LL
		NNSs	NSs	NNSs	NSs	
NP-based	NP with of-phrase fragment (e.g. a large number of the)	40	40	1086(26,5%)	617(27,8%)	130,30***
	NP with other post-modifier fragment (e.g. an important role in)	5	5	133 (10,9%)	71 (10,3%)	19,07***
	SubTotal	45	45	1219	688	149,22***
PP-based	PP phrase with embedded of-phrase (e.g. as a result the)	29	29	1033 (61%)	665(61,9%)	79,97***
	Other PP fragment (e.g. for the first time)	15	15	660 (38,9%)	409(38,1%)	59,20***
	Subtotal	44	44	1693	1074	138,94***
VP-based	Copula be + noun phrase/adjective phrase (e.g. is a part of)	3	2	118 (25,9%)	22 (14%)	72,20***
	VP with active verb (e.g. play a role in)	5	5	53 (11,7%)	42 (26,7%)	1,26
	VP with passive verb (e.g. is based on a)	14	11	283(62,3%)	93 (59,2%)	100,36***
	Subtotal	22	18	454	157	150,33***
Clause-based	PP+copula be (e.g. of this study is)	3	3	121 (19,1%)	25 (16%)	68,60***
	NP+copula be (e.g. results are as follows)	3	3	71 (11,2%)	16 (10,2%)	37,50***
	Anticipatory it+copula be+adjective phrase (e.g. it is necessary to)	9	9	221 (34,9%)	102(65,4%)	44,75***
	Anticipatory it + passive verb+that (e.g. it was found that)	2	2	105 (16,6%)	4 (2,5%)	116,70***
	NP + active verb + that (e.g. results show that the)	6	2	114(18%)	9 (5,7%)	106,00***
	Subtotal	23	19	632	156	307,67***
Conjunctions	Conjunction (e.g. as well as the)	6	6	97	143	8,92*
Total		140	132	4095	2218	564,44***

\*significant at  $p < .01$ ; \*\* significant at  $p < .001$ ; \*\*\* significant at  $p < .0001$ .

**Table 3.** The Proportion of Structural Lexical Bundles

Category	Types		Tokens	
	NNSs	NSs	NNSs	NSs
NP-based	32,1 %	34,1 %	29,7 %	31,1 %
PP-based	31,5 %	33,3 %	41,3 %	48,4 %
VP-based	15,7 %	13,6 %	11,1 %	7,1 %
Clause-based	16,4 %	14,4 %	15,5 %	7,0 %
Conjunctions	4,3 %	4,6 %	2,4 %	6,4 %
Total	100 %	100 %	100 %	100 %

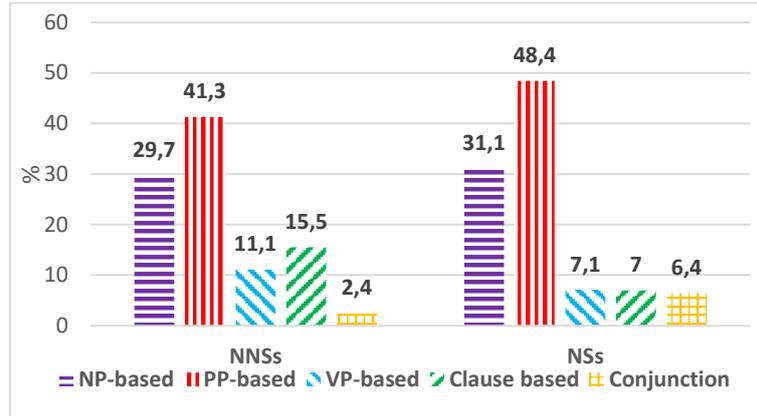


Figure 1. The Proportion of Lexical Bundles Structural Categories

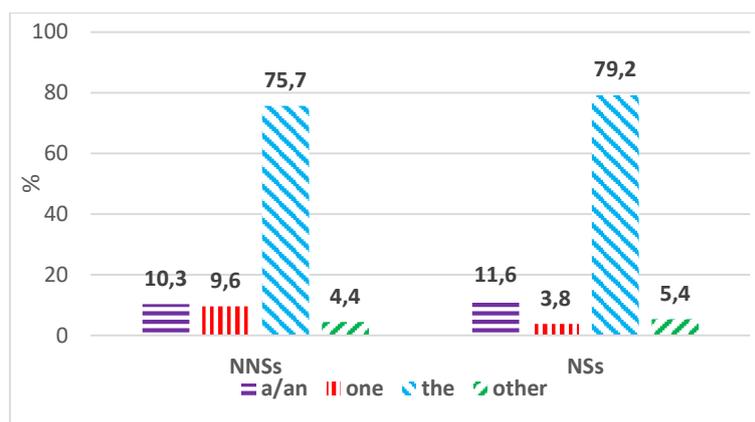
3.2.1. NP-Based Lexical Bundles

NP-based LBs have been re-examined by dividing them into subcategories such as NPs beginning with *a / an, one, the,* and *other NPs*. In general terms, as shown in the table 4., NSSs can be seen that they have used more NP-based LBs than NSs. However, when the results are evaluated in the usage of total lexical bundles and analysed as the proportional level of usage, it can be seen that NSs used more NP-based LBs than NNSs (see Figure 2). In the related literature, it was shown that such NP-based LBs are communication and speech-based structures and more specific to language learners and their writing (Bychkovska & Lee, 2017; Chen & Baker, 2016). However, in this study, it is seen that NNSs have used these structures less than NSs and these structures have shown that they are not only specific to NNSs’ writing.

Table 4. The Subcategory of NP-Based Lexical Bundles

Subcategory	Types		Tokens		LL
	NNSs	NSs	NNSs	NSs	
NP beginning with a/an (e.g. an important role in)	7	7	125	80	9,91*
NP beginning with one (e.g. one of the main)	2	2	117	26	62,53***
NP beginning with the (e.g. the effect of the)	30	30	923	545	98,00***
Other NP (e.g. part of this study)	6	6	54	37	3,17
Total	45	45	1219	688	149,22***

In the analysis of subcategory and proportion level of NP beginning with *a/an* (i.e. a large number of, an analysis of the etc...) it can be easily seen that NSs used more NP-based LBs than NNSs (NSS:29,7%, NSs: 31,1 %). On the other hand, when analysed NP beginning with *one* such as *one of the main* (NNSs: 22, NSs: 6) or *one of the most* (NNSs: 95, NSs: 20), it is concluded that NNSs used a great number of this type of bundle. NP beginning with *the* (i.e. the purpose of this, the result of the, the rest of the etc...) and the *other NP* (i.e. part of this study) based LBs are encountered more in the articles of NSs.



**Figure 2.** The Proportion of NP-Based Lexical Bundles

### 3.2.2. Preposition Based Lexical Bundles

PP-based LBs have been subcategorized as PP-based bundles beginning with *as*, *for*, *in*, *on*, *with* and *other PP*-based LBs. As seen in Table 5, NNSs used significantly more PP-based LBs than NSs. These bundles have been identified as learner bundles in the literature and some researchers have reported that PP-based bundles as phrasal bundles -like NP-based ones-, have been encountered more often in L2 academic writing; however less in native English academic writing (Chen & Baker, 2010; Shin, 2019). However, the present study showed that NSs intended to use these structures more often than NNSs except from the PP-based bundles beginning with *as*, *on* and *with*. However, PP-based bundles beginning with *for*, *in* and *other PP*-based bundles show that NSs tend to use these structures more than NNSs. As an example, some structures such as *at the same time* (NNSs: 49, NSs: 155), *from the perspective of* (NNSs: 17, NSs:25) or *through the use of* (NNSs: 22, NSs: 9) are more idiomatic expressions and preferred more by native writers. A notable point is that, NNSs avoid using first person while NSs tend to use them more [e.g. *in this study I* (NNSs: 5, NSs: 10) or *in this study we* (NNSs: 11, NSs: 27)].

**Table 5.** The Subcategory of PP-Based Lexical Bundles

Subcategory	Types		Tokens		LL
	NNSs	NSs	NNSs	NSs	
PP-based bundles beginning with <i>as</i> (e.g. <i>as a result</i> the)	4	4	166	59	52,88***
PP-based bundles beginning with <i>for</i> (e.g. <i>for a variety</i> of)	4	4	74	53	3,46
PP-based bundles beginning with <i>in</i> (e.g. <i>in the process</i> of)	24	24	715	551	21,12***
PP-based bundles beginning with <i>on</i> (e.g. <i>on the order</i> of)	3	3	431	128	173,06***
PP-based bundles beginning with <i>with</i> (e.g. <i>with respect</i> to the)	2	2	118	32	52,34***
Other PP-based bundles (e.g. <i>through the use</i> of)	7	7	189	251	8,84*
Total	44	44	1693	1074	138,94***

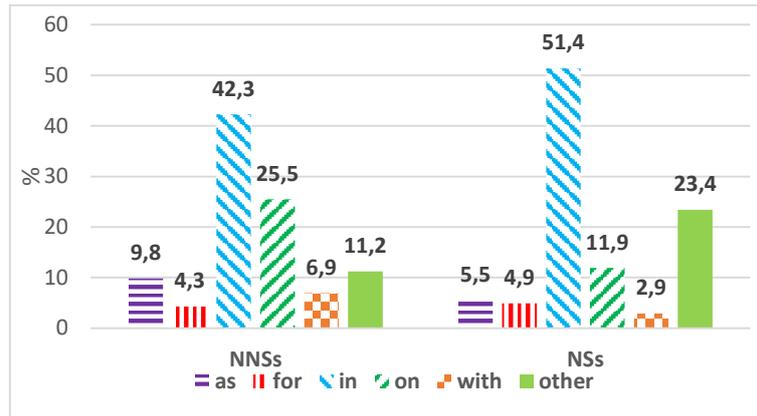


Figure 3. The Proportion of PP-Based Lexical Bundles

3.2.3. Verb Phrase Based Lexical Bundles

VP-based LBs have been divided into three as Copula be + noun phrase/adjective phrase (e.g. is a part of), Verb phrase with active verb (e.g. play a role in) and passive verb (e.g. is based on a). In a consistent with previous studies (Chen& Baker, 2010; Lu & Deng, 2019), NNSs used more VP-based LBs than NSs. Moreover, a particular group of LBs such as *was found that the, were found to be, is found to be* are only found in NNSs’ corpus. Meanwhile, some structures such as *is one of the* (NNSs:91, NSs: 17) or *is a part of* (NNSs 15, NSs: 5) in the group of *Copula be + noun phrase/adjective phrase* are significantly more used by NNSs than NSs. Moreover, the LB *is of great importance* in the same group is only preferred by NNSs (12).

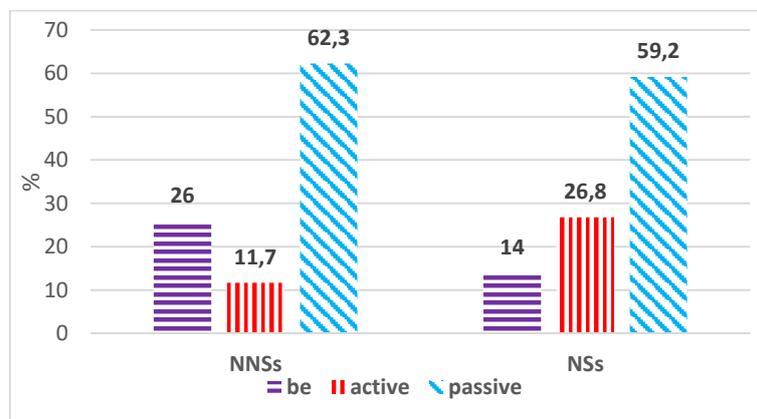


Figure 4. The Proportion of VP-Based Lexical Bundles

3.2.4. Clause Based Lexical Bundles

Figure 5 demonstrates the proportion of LBs usage by both groups. LBs in the group of *PP+ copula be* are preferred frequently by NNSs. In this group, LBs such as *of this study is/are/was* (NNSs: 121, NSs:25) have been used significantly more by NNSs instead of NSs. Similarly, *noun phrase with copula be* (e.g. this study is the/is to/was to) has been used frequently by NNSs. However, *anticipatory it+copula be+adjective* (e.g. it is easier to/it is hard to/it is obvious that etc.) has been preferred more by NSs. *Anticipatory it + passive verb + that* (e.g. it is/was found that) structures within the same group were also largely used by NNSs. Although all papers written in both groups were research articles, NSs did not prefer to use *is found* and *was found* structures. Interestingly, *noun phrase+active verb+that* (e.g. results indicated that) structures have been substantially applied by NNSs. Moreover, NSs did not even choose to use the LBs such as *the results indicated/showed that* although NNSs used them quite a lot. As stated in Biber and Gray (2016) *V+that* structures such as *showed that, indicated that* are more

characteristic of oral speech than academic writing. NNSs of academic writing may not have enough awareness of bundle-genre association which can also be seen the similar findings in the study of Lu and Deng (2019). In this particular study, Chinese students have used more *V+that* structure than English students and the researchers indicate that Chinese students' over-reliance on bundles of this structure may show that they have lack of awareness of bundle-genre association (p. 28).

Conjunctions usage as LBs have been encountered more in NSs' academic writing. Especially, *as well as the/to* (NNSs:53, NSs: 126) have been used more by NSs. This can also show that idiomatic expressions (e.g. *as well as the*) are more chosen by NSs. As Chen and Baker (2010) expressed that LBs in the clausal category are based on more speaking skills and the users should have a higher level of competence, thus, this might be the reason of usage idiomatic expressions by NSs more frequently.

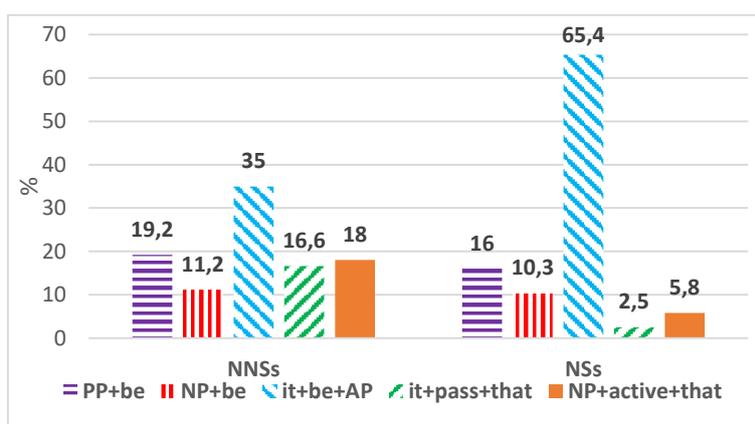


Figure 5. The Proportion of Clause- Based Lexical Bundles

### 3.3. Functional Analysis of Lexical Bundles

To answer the last research question, functional analysis of LBs in terms of Research, Text and Stance have been analysed. Different functional LBs' usage by native and non-native writers have been summarized in Table 6 and the proportion of them have been demonstrated in Table 7 and Figure 5. In terms of different categories and the proportion of them, research-oriented LBs are more frequently used by NSs. On the other hand, text and stance-oriented LBS are more frequently used by NNSs than NSs. This general result shows a parallelism with Lu & Deng's (2019) study in which Chinese students used most frequently text-oriented LBs and research-oriented LBs and finally both groups preferred stance bundles least frequently. Similarly, in the current study, both groups used most frequently text-oriented bundles and least frequently stance bundles. This might be the reason of academic writers' preference of result-oriented bundles, which is in the category of text-oriented bundles.

Table 6. The Categories and Subcategories of Functional Lexical Bundles

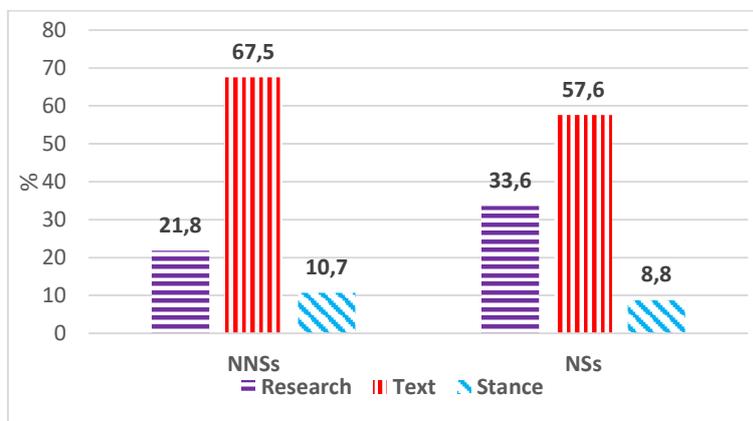
Category	Subcategory	Types		Tokens		LL
		NNSs	NSs	NNSs	NSs	
Research	Location (e.g. in the middle of)	2	2	20 (2,3%)	80 (11,6%)	38,62***
	Procedure (e.g. in the process of)	21	20	403 (45,7%)	323 (46,9%)	8,74***
	Quantification (e.g. is one of the)	6	6	251 (28,5%)	127 (18,5%)	41,30***
	Description (e.g. the nature of the)	6	6	139 (15,8%)	106 (15,4%)	4,42
	Grouping (e.g. a new type of)	5	5	68 (7,7%)	52 (7,6%)	2,12
	Subtotal	40	40	881	688	23,58***
Text	Transition (e.g. as well as the)	6	6	547 (20%)	403 (34,1%)	21,75***
	Comparative (e.g. as well as a)	2	2	18 (0,7%)	15 (1,3%)	0,27
	Resultative (e.g. as a result of)	28	21	1057 (38,7%)	235 (19,9%)	564,69***

Structuring (e.g. goal of this study)	19	19	352 (12,9%)	171 (14,4%)	63,75***	
Framing (e.g. in terms of the)	18	19	629 (23%)	279 (23,6%)	138,06***	
Objective (e.g. to deal with the)	7	7	127 (4,7%)	79 (6,7%)	11,23**	
Subtotal	80	74	2730	1182	627,84***	
Stance	Stance (e.g. it is difficult to)	20	19	430 (10,7%)	181 (8,8%)	104,20***
Total		140	133	4041	2051	659,84***

\*significant at  $p < .01$ ; \*\* significant at  $p < .001$ ; \*\*\* significant at  $p < .0001$ .

**Table 7.** The Proportion of Functional Lexical Bundles

Category	Types		Tokens	
	NNSs	NSs	NNSs	NSs
Research	28,6 %	30,1 %	21,8 %	33,6 %
Text	57,1 %	55,6 %	67,5 %	57,6 %
Stance	14,3 %	14,3 %	10,7 %	8,8 %
Total	100 %	100 %	100 %	100 %



**Figure 6.** The Proportion of Lexical Bundles Functional Categories

3.3.1. Research-oriented Lexical Bundles

When analysed the subcategories of research-based LBs, as stated in Figure 7, NNSs chose to use *quantification, description* and *grouping* LBs whereas NSs preferred more *location* and *procedure* LBs. This result showed a close relationship with the study of Lu and Deng (2019). The researchers revealed that Chinese writers used more research-oriented LBs than English writers. However, the result of current study that NNSs used less location bundle tokens than NSs appeared incongruent with their study’s results since they expressed that Chinese writers used more location bundles than L1 writers did.

When the procedure-oriented LBs analysis is conducted, it can be seen that LBs which was established on the basis of the verb *-to use* was mostly preferred by NSs. For instance, *can be used to* (NNSs:14, NSs: 37), *the use of a/the* (NNSs: 58, NSs:80) or *through the use of* (NNSs: 14, NSs:51) structures are mostly used in NSs’ academic writings.

In the analysis of quantification-oriented LBs, it can be expressed that NNSs mostly preferred numeral LBs such as *a large number of* (NNSs: 21, NSs:15), *as one of the* (NNSs:47, NSs:13), *is one of the* (NNSs: 91, NSs. 17), *the majority of the* (NNSs: 42, NSs: 13). The result that NNSs used more quantification LBs appeared congruent with Lu and Deng’s (2019) findings that Chinese writers used more quantification LBs than L1 English writers, but seemed to be incongruent with Pan, Reppen and Biber’s (2016) findings in which expert writers used more quantification LBs than Chinese writers.

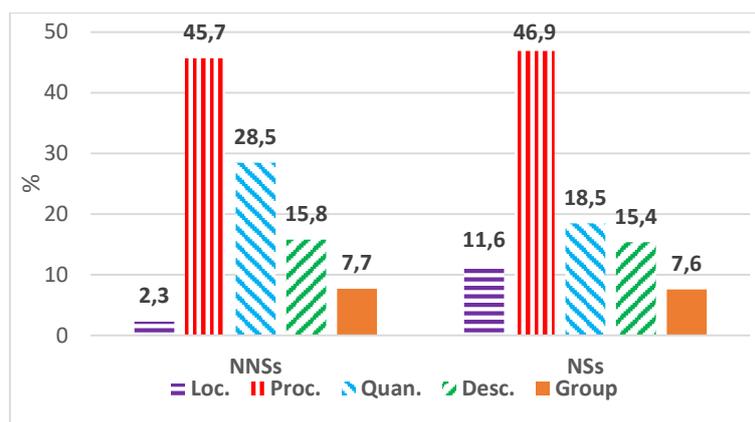


Figure 7. The Proportion of Research-Oriented Lexical Bundles

### 3.3.2. Text Oriented Lexical Bundles

Figure 8 shows the proportion of subcategories in text-oriented LBs. When analysed the *transition bundles*, a significant difference between two groups can be seen and it can be said that the usage of transition bundles is preferred by NSs more frequently than NNSs [e.g. as well as the/to (NNSs: 53, NSs: 126), at the same time (NNSs: 49, NSs: 155)]. Secondly, the *comparative LBs* seem to be used almost equally by both groups. In the third group, it can be seen that the biggest difference between two groups is in the usage of *resultative bundles*. NNSs of academic writing used nearly twice as many resultative bundles as NSs did. To give some examples in this group, the noun form of result has been mostly used by NNSs [(e.g. as a result the/of (NNSs: 116, NSs:37), results indicated/showed that the (NNSs:30, NSs:0), the results indicate/d that (NNSs: 43, NSs:1), the results of the/this (NNSs: 273, NSs:30), the results show/ed that (NNSs: 41, NSs:8)]. The verb form *–to find* has also been most frequently used by NNSs [(e.g. was/were found to be (NNSs: 141, NSs: 1), it is/was found that (NNSs: 105, NSs:4)].

Concerning the next subcategory, it is revealed that *structuring bundles* have been used more frequently by NSs. A detailed examination of the results revealed that NNSs basically used more structuring bundles including the noun *–study* such as *of this study is/are/was* (NNSs: 121, NSs: 25), *this study is/was the/to* (NNSs: 71, NSs: 16). On the other hand, LBs beginning with first person (singular or plural), such as *in this study I/we* (NNSs: 16, NSs: 37) have been mostly used by L1 English writers.

In the context of *framing bundles*, NSs basically depended on PP-based LBs beginning with *in*, [e.g., *in the absence/case/context/face/presence of* (NNSs:97, NSs: 186)], and NP-based bundles including the word *–context*, [e.g. *the context of a/the* (NNSs: 17, NSs: 48)] while NNSs principally tend to use VP-based LBs, [e.g. *is based on a/the* (NNSs: 37, NSs: 14), *on the basis of/the basis of the* (NNSs: 69, NSs: 31)], they also used PP-based LBs in a limited type, [e.g., *in terms of the* (NNSs: 139, NSs:38), *in the field of* (NNSs: 92, NSs: 23)]. As the final group of text-oriented LBs, *objective bundles* have been rarely used by both groups and these bundles have been constructed on the basis of PP-based LBs beginning with *for* and *in*, e.g. *for the purpose of, in order to be/improve*. In their study, Chen and Baker (2010) stated that NNSs tend to use bundles with *in order to* and *to-fragments* more frequently than L1 writers. The results of the present study show a strong relationship to this expression since NNSs used more extensively the bundles such as *in order to be/get//improve* (NNSs: 68, NSs: 18) than NSs.

Also, these results show a close relationship to the findings of Lu and Deng's (2019) findings that Chinese students used more *resultative bundles* than L1 English writers, however, for the other subcategories, the current study appears incongruent with the same study since Chinese students used more bundles in other subcategories such as *framing and objective*.

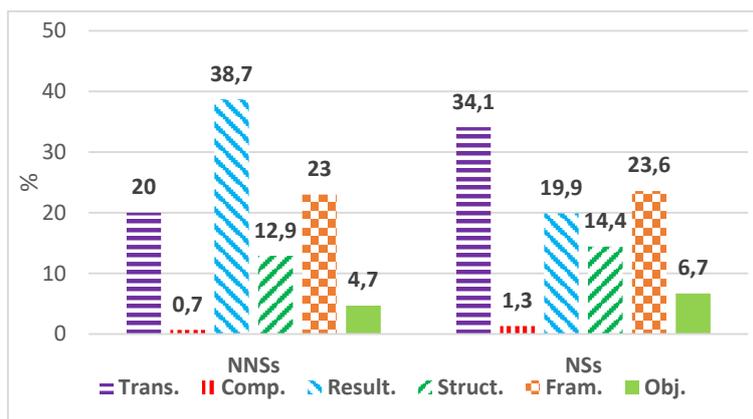


Figure 8. The Proportion of Text-Oriented Lexical Bundles

### 3.3.3. Stance Lexical Bundles

Both groups used least frequent stance LBs in the group of *other functional bundles*. The results getting from stance bundles show that NSs used less stance-oriented bundles than NNSs. A close examination of the data shows that the bundles constructed on the basis of *anticipatory it+be+adj.phrase* were used more frequently by NNSs than NSs, e.g., *it is easier/hard/impossible/difficult/necessary/possible to* (NNSs: 120, NSs: 26). These results are not consistent with Pan, Reppen and Biber's (2016) and Lu and Deng's (2019) findings that non-native students used significantly more stance LBs than native students. In contrast to the results of these studies, the current study showed that NNSs used more stance bundles including the adjective *-important* or the noun *-importance* such as *an important role in, is of great importance, of the most important, play (an) important role(s)* and bundles beginning with *one of the* such as *one of the main/most*.

## 4. Discussion

This study compared the distribution of four-words LBs in different structural and functional categories in academic articles written by native and non-native English writers. There are some similarities and differences between two groups. First, NP-based LBs in structural category were used by both groups in a close proportional level. NP and PP-based LBs in the phrasal bundles group were used more frequently by NSs, whereas VP and Clause based LBs in the clausal category were used more frequently by NNSs. In addition, NSs preferred to use NP-based LBs beginning with *a/an* and *the* more frequently than NNSs. The most notable difference was seen on PP-based bundles beginning with *in* and *on*. In other words, PPs-based bundles beginning with *in* were preferred by NSs, while PPs-based LBs beginning with *on* were preferred by NNSs. In clausal category, VP-based bundles were used more frequently by NNSs and less frequently by NSs. Clause-based LBs were more preferred by NNSs except from *anticipatory it+copula be+adjective* structures. Conjunction-based LBs were less widely used by both groups, but NSs of academic writers still preferred to use them more than NNSs. These results differ from the results of previous literature (Chen & Baker, 2010, Pan, Reppen and Biber, 2016; Shin, 2019). The researchers concluded that LBs in the phrasal category were more frequently used by NNSs, whereas LBs in the clausal category were more frequently used by NSs. They argued that phrasal LBs were more common in academic writing, but that the LBs in the clausal category were based on more speaking skills and the users should have a higher level of competence. However, the results of this study indicate that phrasal bundles are used more than clausal bundles in both groups. It can be said that these results are similar to those of Karabacak and Qin (2013). The researchers examined LBs in English

written argumentative essays obtained from three different groups, L1 Turkish, L1 Chinese and L1 English. According to the results of the comparative analysis, even individuals with advanced English may not be able to fully comprehend these LBs, thus, it was concluded that an explicit learning focused on LBs was necessary. When the meaningful difference between the users was examined, it was seen that there was a significant difference between two groups. In addition, after a detailed examination, it can be seen that although Turkish writers seemed to have used more LBs, these LBs were always the same and frequent repetitions occurred. At this point, the results of the study are similar to those of Öztürk and Köse (2016). Researchers examined the use of four-word LBs in terms of *frequency, function and structure* in the discipline of foreign language learning of postgraduate students whose native language is Turkish and English. The results of the study showed that Turkish postgraduate students used LBs more frequently in their articles. However, when we looked at the number of LBs used, it was seen that Turkish students used the same LBs quite frequently, ie the diversity was low. When the structural and functional categories were examined, it was observed that there was no significant difference between the groups.

From the functional point of view, it was seen that Turkish authors used LBs more than English authors. In terms of research-oriented LBs, NSs used more LBs than NNSs, whereas L1 Turkish users used all other LBs in the subcategory more frequently. In the functional category, after the examination of text-oriented LBs, NNSs used quantification, description and grouping LBs more frequently, while NSs used location and procedure bundles more often. In addition, NSs used LBs to indicate research *objectives*, while NNSs used more LBs to report *results*. At this point, the study results are very similar to those of Lu and Deng (2019). In the functional area, researchers stated that Chinese students used more results-oriented LBs, and L1 English users used more research-oriented LBs. Similarly, in the current study, the use of LBs in the functional category differs from NSs as a result of Turkish authors' efforts to write result-oriented articles in the academic sense. In short, it can be said that Turkish writers take more results-oriented approach in the academic writing world.

## 5. Conclusion

This study revealed some important evidence that NNSs may have fully acquired with the competence of English in academic writing and prefer to write research articles focused on *results*. Since the difference between two groups in the usage of LBs is in low level in academic writing, it can be stated that they may be called as “English users” rather than NNSs. In addition, in the current study, it is seen that, NNSs avoid using LBs on the basis of some nouns such as *function, effort, or context*, but they tend to use more nouns such as *result, fact, and purpose* and fall into repetition. This may be the outcome of NNSs' article writing styles in which they desire to reach meaningful facts or results. To note that, from pedagogical perspective, the results of this study may serve as a tool to teach native-like LBs in the group of multi-word sequences. At university level, the avoidance of falling into repetition and the usage of diversity-oriented multi-word units may be emphasized. The last point as a keynote is that NNSs use the structures beginning with *-one* instead of *a/an or the*. This may be the reason of lexico-grammatical difference between two languages that Turkish language does not consist of definite article *-the* and the accusative suffix (-ı, -i, -u, -ü) is used to state definite case. Thus, Turkish writers may abstain from using definite article *-the*.

In terms of the limitations of the study, the most important one is that, although the biographies of all authors in the study have been examined, some authors may have different backgrounds. Namely, although the article was written by a Turkish author, this author may have lived for many years in a country where English is spoken as the mother tongue, or can have acquired enough competence in the

second language like a native-speaker or may be bilingual or multilingual. Another limitation is that only the articles from two journals were used in this research. This may have resulted in a collection of articles in the similar research area and may indicate the continuous repetition of LBs based on a single or restricted research point.

## 5. Ethics Committee Approval

The author(s) confirm(s) that the study does not need ethics committee approval according to the research integrity rules in their country (Date of Confirmation: August 17, 2020).

## References

- Ädel, 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.
- Allan, R. (2016). Lexical bundles in graded readers: To what extent does language restriction affect lexical patterning?. *System*, 59, 61-72.
- Ang, L. H., & Tan, K. H. (2018). Specificity in English for Academic Purposes (EAP): A Corpus analysis of lexical bundles in academic writing. *3L: Language, Linguistics, Literature*, 24(2).
- Ang, L. H., Tan, K. H. & He, M. (2017). A corpus-based collocational analysis of noun premodification types in academic writing. *3L: The Southeast Asian Journal of English Language Studies*, 23(1), 115-131.
- Anthony, L. (2019). AntConc (Version 3.5.8) [Computer Software]. Tokyo, Japan: Waseda University. Available from <http://www.laurenceanthony.net/software>.
- Bal, B. (2010). Analysis of Four-word Lexical Bundles in Published Research Articles Written by Turkish Scholars. Thesis, Georgia State University, 2010. Available from [https://scholarworks.gsu.edu/alesl\\_theses/2/](https://scholarworks.gsu.edu/alesl_theses/2/)
- Biber, D., & Barbieri, F. (2007). Lexical bundles in university spoken and written registers. *English for Specific Purposes*, 26(3), 263-286.
- Biber, D. (2006). Stance in spoken and written university registers. *Journal of English for Academic Purposes*, 5(2), 97-116.
- 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., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Grammar of spoken and written English*. Harlow, England: Pearson Education Limited.
- 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.
- Chen, Y. H., & Baker, P. (2010). Lexical bundles in L1 and L2 academic writing. *Language Learning & Technology*, 14(2), 30-49.
- Chen, Y. H., & Baker, P. (2016). Investigating criterial discourse features across second language development: Lexical bundles in rated learner essays, CEFR B1, B2 and C1. *Applied Linguistics*, 37(6), 849-880.

- Cortes, V. (2004). Lexical bundles in published and student disciplinary writing: Examples from history and biology. *English for Specific Purposes*, 23(4), 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.
- Ghafar Samar, R., Shokrpour, N., & Nasiri, E. (2018). The applicability of teaching lexical bundles on medical students' writing proficiency in an efl context. *Teaching English Language*, 12(2), 27-44.
- Grabowski, Ł. (2015). Keywords and lexical bundles within English pharmaceutical discourse: A corpus-driven description. *English for Specific Purposes*, 38, 23-33.
- 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.
- Huang, K. (2015). More does not mean better: Frequency and accuracy analysis of lexical bundles in Chinese EFL learners' essay writing. *System*, 53, 13-23.
- Hyland, K. (2008a). Academic clusters: Text patterning in published and postgraduate writing. *International Journal of Applied Linguistics*, 18(1), 41-62.
- Hyland, K. (2008b). As can be seen: Lexical bundles and disciplinary variation. *English for Specific Purposes*, 27(1), 4-21.
- 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.
- 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.
- Meunier, F., & Gouverneur, C. (2007). The treatment of phraseology in ELT textbooks. In *Corpora in the foreign language classroom* (pp. 119-139). Brill Rodopi.
- Nesi, H., & Basturkmen, H. (2006). Lexical bundles and discourse signalling in academic lectures. *International Journal of Corpus Linguistics*, 11(3), 283-304.
- Öztürk, Y., & Köse, G. D. (2016). Turkish and native English academic writers' use of lexical bundles. *Dil ve Dilbilimi Çalışmaları Dergisi*, 12(1), 149-165.
- 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.
- Paquot, M. (2017). L1 frequency in foreign language acquisition: Recurrent word combinations in French and Spanish EFL learner writing. *Second Language Research*, 33(1), 13-32.
- Rayson, P., & Garside, R. (2000, October). Comparing corpora using frequency profiling. In *Proceedings of the workshop on Comparing corpora-Volume 9* (pp. 1-6). Association for Computational Linguistics.

- Ruan, Z. (2017). Lexical bundles in Chinese undergraduate academic writing at an English medium university. *RELC Journal*, 48(3), 327-340.
- Shin, Y. K. (2019). Do native writers always have a head start over nonnative writers? The use of lexical bundles in college students' essays. *Journal of English for Academic Purposes*, 40, 1-14.
- Siyanova-Chanturia, A., Conklin, K., & Van Heuven, W. J. (2011). Seeing a phrase “time and again” matters: The role of phrasal frequency in the processing of multiword sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37(3), 776.
- Staples, S., & Reppen, R. (2016). Understanding first-year L2 writing: A lexico-grammatical analysis across L1s, genres, and language ratings. *Journal of Second Language Writing*, 32, 17-35.
- Tremblay, A., Derwing, B., Libben, G., & Westbury, C. (2011). Processing advantages of lexical bundles: Evidence from self-paced reading and sentence recall tasks. *Language Learning*, 61(2), 569-613.
- Ucar, S. (2017). A Corpus-Based Study on the Use of Three-Word Lexical Bundles in the Academic Writing by Native English and Turkish Non-Native Writers. *English Language Teaching*, 10(12), 28-36.
- Wright, H. R. (2019). Lexical bundles in stand-alone literature reviews: Sections, frequencies, and functions. *English for Specific Purposes*, 54, 1-14.

## Appendix A. Lexical Bundles List

Lexical Bundles	NNSs	NSs	Structure	Function
a better understanding of	12	11	Noun phrase with of-phrase fragment	Stance
a large number of	21	15	Noun phrase with of-phrase fragment	Quantification
a wide range of	28	22	Noun phrase with of-phrase fragment	Grouping
a wide variety of	13	5	Noun phrase with of-phrase fragment	Grouping
an analysis of the	7	13	Noun phrase with of-phrase fragment	Procedure
an important role in	31	12	Noun phrase with other post-modifier fragm.	Stance
an increase in the	13	2	Noun phrase with other post-modifier fragm.	Procedure
and the effects of	11	1	Conjunction	Resultative
and the relationship between	15	1	Conjunction	Resultative
as a function of	3	9	Prepositional phrase with embedded of-phrase	Framing
as a result the	20	5	Prepositional phrase with embedded of-phrase	Resultative
as a result of	96	32	Prepositional phrase with embedded of-phrase	Resultative
as one of the	47	13	Prepositional phrase with embedded of-phrase	Quantification
as well as a	9	10	Conjunction	Comparative
as well as the	48	95	Conjunction	Transition
as well as to	5	31	Conjunction	Transition
at the same time	49	155	Other prepositional phrase fragment	Transition
be used as a	13	4	Verb phrase with passive verb	Procedure
by means of the	10	2	Prepositional phrase with embedded of-phrase	Framing
can be applied to	5	5	Verb phrase with passive verb	Procedure
can be used as	13	4	Verb phrase with passive verb	Procedure
can be used in	13	9	Verb phrase with passive verb	Procedure
can be used to	14	37	Verb phrase with passive verb	Procedure
due to the fact	45	5	Other Prepositional Phrase Fragment	stance
first part of the	9	11	Noun phrase with of-phrase fragment	Structuring
for the development of	14	7	Prepositional phrase with embedded of-phrase	Objective
for the first time	19	6	Other prepositional phrase fragment	Framing
for the purpose of	37	34	Prepositional phrase with embedded of-phrase	Objective
for the study of	4	6	Prepositional phrase with embedded of-phrase	Objective
from the perspective of	17	25	Prepositional phrase with embedded of-phrase	Framing

has been shown to	3	10	Verb phrase with passive verb	Resultative
has the potential to	6	12	Verb phrase with active verb	Stance
have been carried out	10	2	Verb phrase with passive verb	Procedure
have been shown to	3	7	Verb phrase with passive verb	Resultative
have the potential to	7	15	Verb phrase with active verb	Stance
important role in the	7	4	Noun phrase with other post-modifier fragm.	Stance
in a variety of	5	17	Prepositional phrase with embedded of-phrase	Grouping
in addition to the	64	18	Prepositional phrase with embedded of-phrase	Transition
in an effort to	4	14	Prepositional phrase with embedded of-phrase	Objective
in order to be	40	10	Other prepositional phrase fragment	Objective
in order to get	8	6	Other prepositional phrase fragment	Objective
in order to improve	20	2	Other prepositional phrase fragment	Objective
in terms of the	139	38	Prepositional phrase with embedded of-phrase	Framing
in the absence of	2	14	Prepositional phrase with embedded of-phrase	Framing
in the case of	46	52	Prepositional phrase with embedded of-phrase	Framing
in the context of	43	109	Prepositional phrase with embedded of-phrase	Framing
in the development of	19	10	Prepositional phrase with embedded of-phrase	Procedure
in the face of	6	11	Prepositional phrase with embedded of-phrase	Framing
in the field of	92	23	Prepositional phrase with embedded of-phrase	Framing
in the first part	11	4	Other prepositional phrase fragment	Structuring
in the form of	53	48	Prepositional phrase with embedded of-phrase	Description
in the long run	10	2	Other prepositional phrase fragment	Structuring
in the middle of	14	20	Prepositional phrase with embedded of-phrase	Location
in the presence of	11	12	Prepositional phrase with embedded of-phrase	Framing
in the process of	63	32	Prepositional phrase with embedded of-phrase	Procedure
in the second part	10	2	Other prepositional phrase fragment	Structuring
in the study of	33	10	Prepositional phrase with embedded of-phrase	Procedure
in the united states	6	60	Other prepositional phrase fragment	Location
in this study I	5	10	Other prepositional phrase fragment	Structuring
in this study we	11	27	Other prepositional phrase fragment	Structuring
it is easier to	8	2	Anticipatory it + copula be + adjective phrase	Stance
it is hard to	9	1	Anticipatory it + copula be + adjective phrase	Stance
it is impossible to	5	1	Anticipatory it + copula be + adjective phrase	Stance
it is obvious that	30	1	Anticipatory it + copula be + adjective phrase	Resultative
it is true that	6	5	Anticipatory it + copula be + adjective phrase	Resultative
is a part of	15	5	Copula be + noun phrase/adjective phrase	Grouping
is based on a	13	7	Verb phrase with passive verb	Framing
is based on the	24	7	Verb phrase with passive verb	Framing
is found to be	11	0	Verb phrase with passive verb	Resultative
is of great importance	12	0	Copula be + noun phrase/adjective phrase	Stance
is one of the	91	17	Copula be + noun phrase/adjective phrase	Quantification
it is difficult to	19	5	Anticipatory it + copula be + adjective phrase	Stance
it is found that	9	2	Anticipatory it + passive verb + that	Resultative
it is important to	56	70	Anticipatory it + copula be + adjective phrase	Stance
it is necessary to	30	7	Anticipatory it + copula be + adjective phrase	Stance
it is possible to	49	10	Anticipatory it + copula be + adjective phrase	Stance
it was found that	96	2	Anticipatory it + passive verb + that	Resultative
most of the time	19	5	Noun phrase with of-phrase fragment	Structuring
of the most important	32	4	Prepositional phrase with embedded of-phrase	Stance
of this study are	15	2	Prepositional phrase + copula be	Structuring
of this study is	73	15	Prepositional phrase + copula be	Structuring
of this study was	33	8	Prepositional phrase + copula be	Structuring
on the basis of	50	24	Prepositional phrase with embedded of-phrase	Framing
on the one hand	7	20	Other prepositional phrase fragment	Transition
on the other hand	374	84	Other prepositional phrase fragment	Transition
one of the main	22	6	Noun phrase with of-phrase fragment	Stance
one of the most	95	20	Noun phrase with of-phrase fragment	Stance
our understanding of the	1	9	Noun phrase with of-phrase fragment	Stance
part of the study	9	2	Noun phrase with of-phrase fragment	Structuring
play a role in	11	8	Verb phrase with active verb	Resultative

play an important role	20	6	Verb phrase with active verb	Stance
plays an important role	9	1	Verb phrase with active verb	Stance
results indicated that the	15	0	Noun phrase + active verb + that	Resultative
results showed that the	15	0	Noun phrase + active verb + that	Resultative
second part of the	9	6	Noun phrase with of-phrase fragment	Structuring
than that of the	9	5	Conjunction	Comparative
the analysis of the	55	24	Noun phrase with of-phrase fragment	Procedure
the application of the	13	0	Noun phrase with of-phrase fragment	Procedure
the basis of the	19	7	Noun phrase with of-phrase fragment	Framing
the characteristics of the	7	3	Noun phrase with of-phrase fragment	Grouping
the context of a	0	13	Noun phrase with of-phrase fragment	Framing
the context of the	17	35	Noun phrase with of-phrase fragment	Framing
the design of the	9	10	Noun phrase with of-phrase fragment	Procedure
the development of a	6	5	Noun phrase with of-phrase fragment	Procedure
the development of the	16	12	Noun phrase with of-phrase fragment	Procedure
the effect of the	22	3	Noun phrase with of-phrase fragment	Resultative
the effectiveness of the	23	4	Noun phrase with of-phrase fragment	Description
the effects of the	20	1	Noun phrase with of-phrase fragment	Resultative
the end of the	90	69	Noun phrase with of-phrase fragment	Resultative
the extent to which	38	35	Noun phrase with other post-modifier fragm.	Quantification
the first part of	14	14	Noun phrase with of-phrase fragment	Structuring
the focus of this	16	14	Noun phrase with of-phrase fragment	Structuring
the majority of the	42	13	Noun phrase with of-phrase fragment	Quantification
the nature of the	25	41	Noun phrase with of-phrase fragment	Description
the performance of the	11	2	Noun phrase with of-phrase fragment	Description
the purpose of this	26	26	Noun phrase with of-phrase fragment	Structuring
the relationship between the	44	18	Noun phrase with other post-modifier fragm.	Resultative
the rest of the	12	34	Noun phrase with of-phrase fragment	Quantification
the results indicate that	13	1	Noun phrase + active verb + that	Resultative
the results indicated that	30	0	Noun phrase + active verb + that	Resultative
the results of the	204	18	Noun phrase with of-phrase fragment	Resultative
the results of this	69	12	Noun phrase with of-phrase fragment	Resultative
the results show that	9	8	Noun phrase + active verb + that	Resultative
the results showed that	32	0	Noun phrase + active verb + that	Resultative
the role of the	22	30	Noun phrase with of-phrase fragment	Resultative
the second part of	11	7	Noun phrase with of-phrase fragment	Structuring
the structure of the	10	8	Noun phrase with of-phrase fragment	Description
the study of the	7	4	Noun phrase with of-phrase fragment	Procedure
the use of a	11	21	Noun phrase with of-phrase fragment	Procedure
the use of the	47	59	Noun phrase with of-phrase fragment	Procedure
the validity of the	17	3	Noun phrase with of-phrase fragment	Description
this study is the	11	2	Noun phrase + copula be	Structuring
this study is to	46	6	Noun phrase + copula be	Structuring
this study was to	14	8	Noun phrase + copula be	Structuring
through the use of	14	51	Prepositional phrase with embedded of-phrase	Procedure
to the development of	22	9	Prepositional phrase with embedded of-phrase	Procedure
was found that the	20	0	Verb phrase with passive verb	Resultative
was found to be	90	1	Verb phrase with passive verb	Resultative
were found to be	51	0	Verb phrase with passive verb	Resultative
with respect to the	45	16	Other prepositional phrase fragment	Framing
with the help of	73	16	Prepositional phrase with embedded of-phrase	Framing

## Anadili İngilizce olarak yazılmış ve yazılmamış akademik yazında sözcük öbeklerinin kullanımına yönelik bibliyometrik bir çalışma

### Öz

Son yıllarda, akademik dünyada tutunabilmek için akademik araştırmalar yayınlamak zorunlu hale gelmiştir. Akademik yazma becerisinde, sözcüksel demetlerin (lexical bundles) kullanımı, iyi düzenlenmiş paragraflar veya denemeler oluşturmada ustaca bir yetenek olarak vurgulanmaktadır ve İngilizce kullanımında daha doğal görünmek ve İngilizceyi akıcı hale getirmek için gerekli olan dil araçlarıdır. Bu çalışma, ana dili İngilizce olan ve olmayan akademik yazarlar tarafından kullanılan sözcüksel demetlerin fonksiyonlarını, yapılarını ve sıklıklarını karşılaştırmayı amaçlamaktadır. Çalışmada, anadili İngilizce olan ve olmayan akademik yazarlar tarafından yazılan 257 makaleden oluşan bir bütüncü düzenlenmiş ve içerisindeki çok sözcüklü demetler AntConc 3.2.4w programı kullanılarak tanımlanmıştır. Yazılı metinler, anadili İngilizce olmayan konuşmacılar tarafından yazılarak “Journal of Language and Linguistic Studies (JLLS)” dergisinde yayınlanmış makalelerden ve anadili İngilizce olan konuşmacılar tarafından yazılarak “Linguistics and Education Journal (LINGED)” dergisinde yayınlanmış makalelerden toplanmıştır. Çalışmanın bulguları, genel anlamda anadili İngilizce olmayan yazarların daha fazla sözcüksel demet kullandıklarını ancak daha fazla tekrara düştüklerini göstermektedir. Ayrıca, çalışma sonuçları, sözcüksel demetlerin yapısal ve fonksiyonel kullanım sıklıklarında bazı önemli farklılıklar olduğunu göstermektedir. Yapısal kategoride, anadili İngilizce olmayan konuşmacılar ad öbeği (NP) ve ilgeç öbeği (PP) tabanlı sözcüksel demetleri ve bağlaçları anadili İngilizce olan konuşmacılara göre daha düşük sıklıkta kullanırken, eylem öbeği (VP) ve tümcecik tabanlı sözcüksel demetleri daha yüksek sıklıkta kullanmışlardır. İşlevsel kategori açısından, araştırma odaklı sözcüksel demetler, anadili İngilizce olan konuşmacılar tarafından daha sık kullanılırken, metin ve durum odaklı sözcüksel demetler anadili İngilizce olmayan konuşmacılar tarafından daha sık kullanılmıştır. Sonuç olarak, İngilizce ve Türkçe dilleri arasındaki sözdizimsel ve dilbilgisel farklılık ve anadili İngilizce olmayan konuşmacıların sonuç odaklı makale yazma eğilimleri, sözcüksel demetlerin kullanımının belirgin özelliklerini ortaya çıkarmaktadır.

*Anahtar Sözcükler:* sözcük öbeği; akademik yazı; çok sözcüklü yapılar; sözcük öğretimi; İngilizce dili

### AUTHOR BIODATA

Fatma Demiray Akbulut is Asst. Prof. Dr. in Translation and Interpreting Studies Department at Bolu Abant İzzet Baysal University, Bolu. She completed her bachelor and master degree at Gazi University, ELT Department and doctoral degree at Ankara University, Foreign Language Teaching Department. Her research area is Morphology, Morphological Awareness, Lexical Approach in ELT, Bilingualism, Multilingualism and Sociolinguistics. She has been currently researching English language teaching from multidisciplinary perspectives.