Derlem Tabanlı Bir Çalışma: Türk Öğrencilerinin Akademik Yazılarında Sözcük Öbekleri Kullanımı

Serpil UÇAR[1]

Elham ZARFSAZ[2]

Öz

Bu çalışmanın amacı, lisans öğrencilerinin akademik yazılarındaki sözcük öbeklerini yapısal ve işlevsel olarak belirlemektir. Araştırmanın derlemi, İngiliz Dili Eğitimi ve İngiliz Dili ve Edebiyatı bölümlerindeki 80 Türk ve 50 İngiliz lisans öğrencilerinin tartışmacı makalelerinden derlenmiştir. Sonuçlar, Türk lisans öğrencilerinin anadili İngilizce olan öğrencilere göre daha az çeşitli ve daha sınırlı sayıda sözcük öbeği kullandığını göstermiştir. Yapısal sınıflandırma bakımından, anadili İngilizce olan öğrencilerin akademik yazılarında daha fazla –of- parçası içeren isim sözcük öbekleri kullanılmıştır. Benzer şekilde, Türk lisans öğrencileri de aynı kategoride sözcük öbeği kullanma eğiliminde olmuşlardır. İşlevsel sınıflandırma ile ilgili olarak, ana dili İngilizce olan öğrenciler, ana dili İngilizce olmayan öğrencilere göre çok daha fazla gönderimsel ve tutum sözcük öbekleri kullanmışlardır. Diğer taraftan, Türk öğrenciler ise söylem belirleyici sözcük öbeklerini daha çok kullanmışlardır.

Anahtar Kelimeler: Derlem tabanlı çalışma, sözcük öbekleri, akademik yazma

A Corpus-Based Study: The Employment of Lexical Bundles in Turkish Students' Academic Writing

Abstract

The current study investigated the use of lexical bundles functionally and structurally in the argumentative papers of undergraduate students. The corpora were 80 argumentative papers of Turkish non-native undergraduate students English Language Teaching department at Gaziosmanpaşa University in Tokat, Turkey and 50 argumentative papers of English native undergraduate students of the English Language and Literature department in the (BAWE) Corpus. 20 times occurrences and appearance in 5 different texts reflected the description of a lexical bundle in the study. The findings revealed that Turkish university students had less diversified and more restricted bundles in number than English native students. In terms of structural classification, more NP phrases with of-phrase fragments were employed by native and non-native student academic writing. In terms of the taxonomy of functions, native students had a greater proportion of referential and stance bundles although non-native students employed more discourse organizers than their native counterparts.

---

[1] Sorumlu Yazar : Serpil UÇAR, Dr. Öğr. Üyesi,Tokat Gaziosmanpaşa Üniversitesi,Türkiye, serpil.ucar@gop.edu.tr ,0000-0001-9722-8236

[2] Elham ZARFSAZ, Dr. Öğr. Üyesi, Tokat Gaziosmanpaşa Üniversitesi, Türkiye, elham.zarfsaz@gop.edu.tr, 0000-0002-6172-6388

Introduction

Recurrent word combinations in writing have drawn incremental attention in recent studies over the past years (Biber, 1999; Jones & Haywood, 2004; Cortes, 2004, 2006; Allen, 2009). Altenberg (1998) stated that 80% of the words constitute recurring word combinations in London-Lund Corpus whereas word combinations construct over 50 percent both of spoken corpus (58, 6%) and written discourse (52,3%) (Erman & Warren, 2000).

According to corpus-based research, learners' usage of lexical bundles is often troublesome (Allen, 2009; Juknevičienė, 2009; Salazar, 2010; Adel & Erman, 2012). For example, Juknevičienė, (2009) showed that L2 learners underused NP with post-modifier fragments. Li & Schmitt (2009) argued that learners employ a variety of native-like lexical bundles, however, they tended to overuse such bundles through which their writing was considered as non-native. Allen (2009) found that Japanese learners of English overused the *-NP + of-* constructions. Adel & Erman (2012) stated that "non-native speakers exhibit a more restricted repertoire of recurrent word combinations than native speakers". Moreover, Chen and Baker (2010) assumed that learners underused referential lexical bundles and overused certain bundles.

A great deal of research was dedicated to corpus-based research emphasizing the usage of formulaic phrases in written discourse and conversation (Cortes, 2004; Biber & Barbieri, 2007), the usage of lexical bundles between native and learner writings (Wei & Lei, 2011; Öztürk, 2014) or on the instructional aspect of these combinations (Li & Schmitt, 2009; Alhassan & Wood, 2015). However, to researcher knowledge, there is not much research about the lexical bundle usage in British and Turkish undergraduate students' academic writing in different disciplines (Adel & Erman, 2012, Cortes, 2004). Therefore, this current study investigated to reveal lexical bundles functionally and structurally in academic papers of English-native and Turkish non-native undergraduate (freshman and sophomore) students in the disciplines of English Language Literature and English Language Teaching.

Literature Review

There are a variety of terminologies for formulaic sequences such as "prefabricated patterns" (or prefabs) (Granger, 1998); "clusters" (Scott, 1999); "fixed expressions" and "idioms" (Moon, 1998); "collocations" (Sinclair, 1991; Howarth, 1998b; Gitsaki, 1999); "recurring word combinations" (Butler, 1997). All the research identifies formulaic sequences in different definitions including a variety of measures and characterizations.

Biber et al. (1999) first described the concept of "lexical bundle" in the Longman Grammar of Spoken and Written English comparing the most frequently used lexical bundles in conversation and written registers. In this book, the lexical bundle was defined as "recurrent expressions, regardless of their idiomaticity, and regardless of their structural status", basic phrases of three or more words that are mostly used together within spoken and written registers. Cortes (2004) also defined them as "extended collocations of three or more words that statistically co-occur in a register." Another

definition was "the most frequent recurring lexical sequences ..., which can be regarded as extended collocations: sequences of three or more words that show a statistical tendency co-occur." (Biber & Conrad, 1999)

Allen (2009) analyzed the frequency and the type of lexical bundles engaged by undergraduate Science Students in academic writing in the ALESS Corpus at the University of Tokyo. Each writing contains the sections of abstract, introduction, method, results, discussion, and conclusion. The results demonstrated that a significant difference emerged between lexical bundles employed by learners and native professional authors. As for the structures of lexical bundles, the learners overused noun phrase constructions, "-*NP + of*-".

Another study conducted by Chen & Baker (2010) investigated the lexical bundle usage in English and Chinese learners' writing to reveal the possible troublesome areas in second language acquisition. The learner corpus constitutes Chinese learners' academic writing although the other two corpora include writing from native scholars and undergraduate students. The academic writing of native and non-native students showed similarity, in which discourse markers and VP-based bundles appeared more in non-native academic writing whereas native students used NP-based bundles and referential bundles more in their academic writing.

Likewise, Adel & Erman (2012) had an investigation on lexical bundles engaged by Swedish and British native university students in the Linguistic department. The corpus, which included 325 essays was Stockholm University Student English Corpus (SUSEC). The corpus was gathered from Swedish and British undergraduate students. The findings demonstrated that L1 speakers of Swedish had a tendency to make use of more restricted and less diversified lexical bundles. More diversified lexical bundles utilized by native speakers were "in unattended 'this' constructions, existential 'there' constructions, hedges and passive constructions" (Adel & Erman, 2012).

Karabacak & Qin (2013) examined the lexical bundle occurrence in the argument essays of three groups of undergraduate students; Chinese, Americans, and Turkish. The corpus was formed from the argumentative articles of Turkish and Chinese sophomores and American freshmen as the learning corpus, and New York Times articles were used as the source corpus. The results demonstrated that even advanced English learners had trouble with the acquisition of certain lexical bundles through limited exposure.

In light of these studies, although there have been quite a number of studies about the lexical bundles in learner writing, there needs to have more empirical studies examining lexical bundle usage in Turkish undergraduate students' writing in different disciplines. Accordingly, to achieve its aims, the study addressed the research questions below:

RQ1: What are the most frequent lexical bundles in the argumentative papers of undergraduate students?

RQ2: What are the structural taxonomies of lexical bundles in the argumentative papers of undergraduate students?

RQ3: What are the functional taxonomies of lexical bundles in the argumentative papers of undergraduate students?

Method

This part consists of the characteristics of the corpus material and analytical steps to be taken in the process of analysis.

Research Corpora

One of the research corpora is from argumentative essays of Turkish freshman and sophomore students at ELT at Gaziosmanpaşa University in Tokat, Turkey. The non-native material consists of 80 final argument-essays collected over two semesters. Each paper has a thesis statement, introduction, body paragraphs, and conclusion. The students were at the same academic level (1st year) and age (18-20).

The other material used for the study includes 50 essays of native English freshman and sophomore undergraduate students in the discipline of ELL in the British Academic Written English Corpus (hereafter BAWE). BAWE is made up of academic writing at universities in the UK. It represents a pattern of British academic English including a variety of disciplines such as Arts and Humanities, Life Sciences. The BAWE corpus includes approximately 3000 texts of British student writing including a variety of genres totaling 6.5 million words (Allen, 2009). The details about the research corpora were given in Table 1.

Table 1. The details about research corpora

	Student Level	Number of Essays	Corpus size (word count)	Average Essay Length
Native students	Freshman /sophomore	50	71.205	600-1400
Non-native students	Freshman/ sophomore	80	52.882	600-1200

Analytical Steps

The study only concentrated on 3-word lexical bundles since they are widely distributed in academic writing in respect to the 4 and 5-word bundles. The perspective developed by the research of Simpson-Vlach and Ellis (2010, p.59) who claimed that a great majority of prominent lexical bundles are 3-word bundles which was adopted by this present study. As Cortes (2004) and Hyland (2008a) set frequency cut-off points at 20 times per-million words, their criteria were applied by the researcher. 20 times occurrences and appearing in 5 different texts reflected the description of a lexical bundle in the current study.

Ant Conc 3.5.9 corpus analysis toolkit was utilized in order to retrieve the bundles within the corpora. This tool listed the lexical bundles with cut-off points of 20 occurrences within 5 different texts. The researcher excluded content bundles including proper names, words related to the topic and disciplines, tables, references, figures, and charts. Each bundle was analyzed qualitatively through the Ant Conc 3.5.9 program to define the types, tokens, and range of the retrieved bundles. The categorization of Biber et al. (1999) was used for the structural classification, the taxonomy of Biber et al.'s (2004) and Simpson-Vlach and Ellis' (2010) Academic Formulas List (AFL) were utilized for the classification of functional categories for the bundles.

Results and Discussion

The Frequency of Bundles

In this part, the tokens and types of bundles in English native and Turkish non-native undergraduate students' academic writing were analyzed in detail. Table 2 listed the total number of lexical bundles in students' writing.

Table 2. Frequency of Lexical bundles

	Bundle Type	Bundle Token	Corpus Size
English Native Students	16	513	71.205
Turkish Non-native Students	7	295	52.882

After the elimination of the unrelated bundles (such as proper names, and context-dependent bundles), as seen in Table 2, English native undergraduate students' academic writing employed a higher number and variety of lexical bundles (n=513/16) compared to Turkish non-native students (n=295/7). Therefore, it can be concluded that Turkish students had less diversified and more restricted quantity of bundles compared to English native students. Moreover, frequencies per-million words and per-texts were also measured to reveal the standardized results of the corpora.

Table 3. Frequencies per million words & texts between the corpora

	Raw Frequency	per million words	per text
Non-native Student Corpus	295	5578	3.68
Native Student Corpus	513	7204	10.26

As seen in Table 3, results showed that Turkish students engaged fewer lexical bundles per-million words and per-texts (5578 occurrences per million words and an average of 3.68 per text) than English native students (7204 occurrences per million words and an average of 10.26 per text). In respect to the lexical bundle range and type, Table 4 shows the types, tokens, and ranges of the retrieved bundles in Turkish and English undergraduate learners' argumentative papers.

Table 4. Type, token and range of lexical bundles

Turkish Students' Bundle Type	Bundle Tokens	Bundle Range	English Students' Bundle Type	Bundle Tokens	Bundle Range
as a result	56	30	"the use of"	85	35
the development of	55	31	"in order to"	39	18
one of the	48	33	"a sense of"	37	21
the growth of	40	16	"the end of"	35	22

in order to	38	22	"the fact that"	32	17
the use of	34	19	"way in which"	32	10
due to the	24	21	"in the second"	31	18
			"in the first"	30	20
			"it is a"	29	21
			"that it is"	29	22
			"there is a"	27	18
			"the idea of"	22	18
			"there is no"	22	19
			"in terms of"	21	12
			"nature of the"	21	15
			"the effect of"	21	11
Total	295			513	

According to Table 4, the most frequent bundle employed by Turkish native learner argumentative papers was "as a result" which was identified 56 times across 30 different texts whereas English native students' academic writing employed the bundle of "the use of" as the most frequent three-word bundle with a frequency of 85 times across 35 different texts. Comparing the two corpora in terms of the most frequent bundles, only two of these bundles were shared bundles which were "in order to" and "the use of". Therefore, it seemed that Turkish and English students used different bundles in their academic writing.

Moreover, even though the total frequency of lexical bundles in Turkish students' writing was much lower than in native student writing, the frequency of 4 bundles was over 40 times while native student writing had only one bundle with the same frequency. However, some bundles were overused by Turkish students. One typical example was the second most frequent bundle "the development of" which occurred 55 times across 31 texts but was not in the bundle list of the English native student corpus. It seemed that Turkish non-native students overused the bundle in their writing.

Structural Taxonomy

The structural classification of the bundles was based on the categories created by Biber et al. (1999). Table 5 demonstrates the lexical bundle structures used in the essays of English and Turkish students.

Table 5. Structures of lexical bundles

Structure	Turkish Student Corpus	English Student Corpus

"Noun phrase with of- phrase fragment"		"the use of"
	"the development of"	"a sense of"
	"the use of"	"the end of"
	"one of the"	"the idea of"
	"the growth of"	"nature of the"
		"the effect of"
"Noun phrase with other post-modifier fragments"		"the fact that"
		"way in which"
"Prepositional phrase with embedded of-phrase fragment"	"as a result (of)"	"in terms of"
"Other PP (fragment)"	"in order to"	"in order to"
	"due to the"	"in the first"
		"in the second"
"Anticipatory it + verb / adjective phrase"		
"Passive verb + prepositional phrase fragment"		
"Copula be + noun / adjective phrase"		
"(Verb phrase+) that- clause fragment"		"that it is"
"(Verb/ adjective +) to-clause fragment"		
"Adverbial clause fragment"		
"Pronoun/ noun phrase+ be (+…)"		"there is a"
		"there is no"
		"it is a"
"Other expressions"		

Biber et al., 1999, p.1014-1024

Table 5 demonstrated that native student academic writing employed more NP + of- phrase fragments (the use of, a sense of, the end of, the idea of, nature of the, the effect of). Similarly, Turkish student writing tended to use more bundles in the same category (the development of, the use of, one of the, the growth of). Three structural types of lexical bundles were both used by the two corpora; NP + of phrase fragment (the development of, the use of, one of the, the growth of, a sense of, the end of, the idea of, nature of the, the effect of), PP with embedded of-phrase fragment (as a result), other PP (due to, in the first, in order to, in the second). One of the results showed that for example the native list had the "existential "there" constructions pattern but this pattern did not appear in the non-native list. According to table 5, it can be concluded that Turkish students had more limited structural categories than English students. Table 6 lists the distribution of structural taxonomies of lexical bundles employed in two corpora.

Table 6. Frequencies and Percentages of Structural Lexical Bundles

	Non-native Students' Corpus		Native Students' Corpus	
	Frequency	Percentage	Frequency	Percentage
"Noun phrase with of- phrase fragment"	177	60%	221	43.05%
"Noun phrase with other post-modifier fragments"	-	-	64	12.5%
"Prepositional phrase with embedded of-phrase fragment"	56	19%	21	4.10%
"Other prepositional phrase (fragment)"	62	21%	100	19.50%
"(Verb phrase+) that- clause fragment"	-	-	29	5.65%
"Pronoun/ noun phrase+ be (+…)"	-	-	78	15.2%
Total	295	100%	513	100%

As seen in Table 6, while a greater proportion (over 50%) of lexical bundles were based on a single structural category (NP phrase with of- phrase fragment) in the Turkish learner corpus, different proportions of lexical bundles were distributed in six structural categories in native student corpus.

Functional Categories of the Target Bundles in the Corpora

Three prominent functions were as follows: "stance bundles, discourse organizers, and referential bundles" (shown in Table 7*). The definitions of these bundles were as follows (Biber et al., 2004):

"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 the entity or to single out some particular attribute of the entity as especially important."

Table 7 listed the functional distribution of lexical bundles in the corpora.

Table 7. Functions of the Target Bundles in Students' Corpora

Category	Turkish Non-Native Student Corpus	English Native Student Corpus
Stance Bundles	-	"the fact that
Discourse Organizers	"in order to" "due to the" "as a result (of)"	"in order to" "the effect of" "in the first" "in the second"

Referential Bundles	"one of the"	"it is a"
	"the development of"	"that it is"
	"the growth of"	"the way in which"
	"the use of"	"there is a"
		"the use of"
		"there is no"
		"nature of the"
		"in terms of"
		"a sense of"
		"the idea of"
		"the end of"

Biber et al., (2004)

However, the fundamental issue was that there was no obvious criteria for how to decide in which functional class a target bundle should pertain to. For instance, "Focusing is labelled as a discourse organizing in Chen & Baker but referential in Biber et al (2004) and Simpson, Vlach and Ellis (2010)" (Adel & Erman, 2012). Therefore, the researcher made a functional classification considering Biber et al's (2004) research and Simpson, Vlach, and Ellis' (2010) AFL list as references.

In Table 7, referential bundles and other bundle types such as stance bundles appeared in the writing of English native students much more than those of Turkish students' writing. One example of the stance bundle is "the fact that" solely appeared in the native list.

Table 8. Functional Categories of Bundles

	Non-native Students' Corpus		Native Students' Corpus	
	Frequency	Percentage	Frequency	Percentage
Referential Bundles	177	60%	360	70.27%
Discourse Organizers	118	40%	121	23.50%
Stance Bundles	-		32	6.23%
Total	295		513	

Table 8 shows that native students had a greater proportion of referential (70.27%) and stance bundles (6.23%) while non-native students employed more discourse organizers (40%) than their native counterparts as they produced much fewer types of lexical bundles than native students. A great many of lexical bundles appeared in native students' writing did not appear in the non-native texts. For instance, the stance bundle "the fact that" (noun + complement clause combination) was only used by native students.

- "*The fact that* she dismisses much of what the pastor says reflects the gradual dismissal of religion in society at the time"

- The stress on the questioning word emphasizes *the fact that* it is a dilemma; the writer is asking us a question.

The functional category of "the fact that" is the impersonal epistemic stance bundle which functions as expressing degrees of certainty instead of uncertainty (Biber, Conrad & Cortes, 2004).

Another referential bundle that was only used by native students was "(the) way in which" (NP with other post-modifier fragments);

- "The _way in which_ people sought to better themselves individually essentially resulted in a more 'public' sense of the image."

- "This relates to the _way in which_ the Romantics regarded nature as "potentially the landscape of the mind itself"

The referential bundles such as "way in which, nature of the, in terms of, the use of, a sense of, the idea of" were in the category of specification of attributes (intangible framing attributes). This class contains concrete and abstract concepts or categories (Simpson, Vlach & Ellis, 2010). These phrases are clearly important and common academic phrases in written genres.

- "The structure of these end lines gives _a sense of_ finality by summarizing these polarities within two lines"

- "_In terms of_ rhyme scheme, the first and third stanzas consist of three rhyming couplets."

- "The word 'played' enhances _the idea of_ youth and the childish nature of the boy and the foolishness of the situation."

- "Aphra Behn begins this poem by relating the narrator's passion to a stream, through _the use of_ metaphorical language".

On the other hand, Turkish non-native writing employed three referential bundles (one of the, the growth of, the development of) and two discourse organizers (due to the, as a result) which were not used by native counterparts:

- "…animation games teach children to empathize, contribute to _the development of_ creativity and imagination…"

- "Today, there is an inverse relationship between _the growth of_ the economy and the protection of nature…"

- "_As a result_, computers have very significant adverse effects such as anxiety…"

- "...the level of air pollution increases _due to the_ fumes…"

Conclusion

The current study investigated the use of lexical bundles structurally and functionally in the argumentative papers of undergraduate students. The corpora were 80 argumentative papers of Turkish non-native undergraduate students ELT department at Gaziosmanpaşa University in Tokat, Turkey and 50 argumentative papers of English native undergraduate students of ELL in the (BAWE) Corpus. 20 times occurrences and appearance in 5 different texts reflected the description of a lexical bundle in the study.

The results demonstrated that Turkish undergraduate students had less diversified and a more limited number of bundles than English native students. The most frequent bundle employed by Turkish native learner argumentative papers was "as a result" which was identified 56 times across 30 different texts whereas English native students' academic writing employed the bundle of "the use of" as the most frequent three-word bundle with a frequency of 85 times across 35 different

texts. Comparing the two corpora of freshman and sophomore undergraduate student writing in terms of the most frequent bundles, only two of these bundles were shared bundles which were "in order to" and "the use of". Therefore, it seemed that Turkish and English students used different bundles in their academic writing. As for the structural taxonomy, native student academic writing employed more NP phrases with of- phrase fragments (the use of, a sense of, the end of, the idea of, nature of the, the effect of). Similarly, Turkish student writing tended to use more bundles in the same category (the development of, the use of, one of the, the growth of). Three structural types of lexical bundles were both used by the two corpora; NP with of-phrase fragment (the development of, the use of, one of the, the growth of, a sense of, the end of, the idea of, nature of the, the effect of), PP with embedded of-phrase fragment (in terms of, as a result), other PP (due to, in order to, in the first, in the second). In respect to the functional taxonomy, native students had a greater proportion of referential and stance bundles although non-native students employed more discourse organizers than their native counterparts as they produced much fewer types of lexical bundles than native students.

The current study was parallel with the former research that indicated non-native learners used more restricted and less diverse lexical bundles in their academic writing (Cortes, 2004; Allen, 2009; Chen & Baker, 2010). Oktavianti and Sarage (2021) investigated the usage of 3-5 word bundles employed by university-level students in their writing argument essays. The research corpus was obtained from essays of students' as a learner corpus. The findings indicated that NP-based bundles are the most frequently used structural class in the corpus. Moreover, the study also demonstrated that "the use of" was identified as the most frequent individual bundle throughout the research.

Another study conducted by Uçar (2017) investigated the frequencies, structures, and functions of lexical bundles in the published articles of native and non-native scholars. The results showed that Turkish writers utilized less diverse lexical bundles and underused certain bundles in their academic prose. Also, the study showed that the most frequent bundle was "the use of" which was also employed as the most frequent bundle type in native student writing in this current study. Biber (2010) states that a great number of bundles in written prose contain NP with of- phrase fragment (e.g., the use of the) or PP with embedded of-phrase fragment (e.g., as a result of)", which constitutes parallel idea with the findings of the current research.

Pedagogical Implications

Lexical bundles could not be easily acquired in the learning process regarding the findings of the study. One possible solution to this problem could be explicit instruction in which writing instructors should integrate through a group of activities of the target bundles in order to develop learners' academic writing skills.

Computational linguistics has gained popularity as a teaching tool in language teaching. Concordancers could assist reduce the burden on teachers. With thousands of texts put together, concordances can provide data that is fun to work with, especially for higher-level students. It is now widely accepted that "corpus definitions also improve our understanding of fixed units of expression, collocations, and broader language patterns" (O'Keeffe et al., 2007). This new approach, which is referred to as "data-based learning" in the literature, has the ability to "make the student a language researcher" (Johns, 2002). It is also the responsibility of teachers to introduce students such tools by carefully selecting or designing adaptive exercises.

Limitations of the Study

This study has certain limitations that might be addressed in further research. The first restriction was the small size of participants included in this study. More participants can be included to make

broader generalizations to the population. Because of the small size of this specialized corpus, a larger amount of data may provide more reliable, valid, and objective results to investigate LBs as a further study. In the current study, only the argumentative type was analyzed in terms of the use of LBs. Different types of genres could be included to examine students' different genre-specific LBs.

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.

AlHassan, L., & Wood, D. (2015). The effectiveness of focused instruction of formulaic sequences in augmenting L2 learners' academic writing skills: A quantitative research study. *Journal of English for Academic Purposes*, *17*, 51-62.

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

Altenberg, B. (1998). On the phraseology of spoken English: The evidence of recurrent word-combinations.

A.P. Cowie (Ed*.), in *Phraseology: theory, analysis and applications* (pp.101-122). Oxford: Oxford University Press.

Biber, D. (2010). Corpus-based and corpus-driven analyses of language variation and use. In B. Heine & H. Narrog (Eds.), The Oxford handbook of linguistic analysis (pp. 159–191). Oxford: Oxford University Press.

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

Biber, D. & Conrad, S. (1999). Lexical bundles in conversation and academic prose. H. Hasselgard and S. Oksefjell (Eds.), in *Out of corpora* (pp. 181-190). Amsterdam: Rodopi

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). *Longman grammar of spoken and written English*. Harlow, England: Longman

Butler, C. (1997). Repeated word combinations in spoken and written text: some implications for Functional Grammar. C. Butler, J. Connolly, R. Gatward, and M. Wismans (Eds), in *A fund of ideas: Recent development in Functional Grammar* (p. 60-77). Amsterdam: Institute for Functional Research into Language and Language Use.

Chen, Y. H. & Baker, P. (2010). Lexical bundles in L1 and L2 academic writing. *Language Learning & Technology*, 14 (2), 30-49.

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. (2006). Teaching lexical bundles in the disciplines: An example from a writing intensive history class. *Linguistics and Education*, 17 (4), 391-406.

Erman, B. & Warren, B. (2000). The idiom principle and the open choice principle. *Text-Interdisciplinary Journal for the Study of Discourse*, 20 (1), 29-62

Gitsaki, C. (1999). Teaching English collocations to ESL students. *NUCB Journal of Language Culture and Communication*, 1 (3), 27-34.

Granger, S. (1998). Prefabricated patterns in advanced EFL writing: Collocations and formulae. A.P. Cowie (ed.), in *Phraseology: Theory, Analysis and Applications* (pp. 145-160). Oxford: Oxford University Press

Howarth, P. (1998b). The phraseology of learners' academic writing. A.P. Cowie (Ed.), in *Phraseology: theory, analysis and applications* (pp. 161-186). Oxford: Oxford University Press

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

Johns, T. (2002). Data-driven learning: The perpetual challenge. In B. Ketteman, & G. Marko (Eds.), Teaching and Learning by Doing Corpus Linguistics (pp. 107-117). Amsterdam: Rodopi.

Jones, M., & Haywood, S. (2004). Facilitating the acquisition of formulaic sequences. N. Schmitt (Ed.), in *Formulaic sequences acquisition, processing and use*, (pp.269-292). Philadelphia: John Benjamins.

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.

Li, J. & Schmitt, N. (2009). The acquisition of lexical phrases in academic writing: A longitudinal case study. *Journal of Second Language Writing*, 18 (2), 85-102.

Moon, R. (1998). *Fixed expressions and idioms in English: A corpus-based approach*. NewYork: Oxford University Press.

O'Keeffe, A. McCarthy, M., & Carter, R. (2007). From Corpus to Classroom: Language Use and Language Teaching. Cambridge: Cambridge University Press.

Oktavianti, I. N., & Sarage, J. (2021). Lexical Bundles in Students' Argumentative Essays: A Study of Learner Corpus. *Indonesian Journal of EFL and Linguistics*, *6*(2), 509-534.

Öztürk, Y. (2014). *Lexical bundle use of Turkish and native English writers: A corpus-basedd study.* Unpublished MA thesis. Eskişehir: Anadolu University Graduate School of Educational Sciences.

Salazar, D., 2010. Lexical bundles in Philippine and British scientific English. Philipp. J. Linguist. 41, 94--109.

Scott, M. 1999. Wordsmith Tools. Software. Oxford: Oxford University Press.

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

Simpson-Vlach, R. & Ellis, N. C. (2010). An academic formulas list: New methods in phraseology research. *Applied Linguistics*, 31 (4), 487-512.

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.

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