THE JOURNAL OF TEACHING ENGLISH FOR SPECIFIC AND ACADEMIC PURPOSES
Vol. 7, N° 2, 2019, pp. 225–235

UDC: 811.111:338 https://doi.org/10.22190/JTESAP1902225D

STRUCTURE OF LEXICAL BUNDLES IN ECONOMICS RESEARCH ARTICLES

Violeta Damchevska

Centre for Foreign Languages, Ministry of Defence, Macedonia
Phone: +38970670728, E-Mail: v.damchevska@yahoo.com

Abstract. *This paper looks into the structure of four-word lexical bundles in scientific articles in English in economics journals, written by Macedonian scholars. It adopts the lexical bundle method, which requires multi-word sequences to be identified in an electronic corpus, based on their frequency and fixedness. For the purpose of the research, a corpus of economics research articles was compiled, based on which lexical bundles were identified, and compared with bundles found in four other disciplines. The results suggest that academic writers in the discipline of economics rely on the use of lexical bundles, and share bundles found among the twenty most frequent in the academic prose. The bundles identified in published economic writing show similarity with the bundles found in applied linguistics and business studies, and differences with the ones used in biology and electrical engineering. The grammatical structure of the bundles was found to be clausal, with the majority of bundles containing noun phrases and prepositional phrases. The findings from this research have some implications for EAP curricula developers.*

Key words: *lexical bundles, corpus, economics research articles, academic writing*

1. INTRODUCTION

Many linguistic studies of academic discourse in the past few decades have focused on the description of multi-word expressions (eg. Nattinger and DeCarrico, 1992; Lewis, 1993; Schmitt, 2004), challenging the traditional dichotomy between grammar and vocabulary. These studies are part of a long line of research that goes back to Firth (1951), Altenberg (1988), Sinclair (1991), suggesting that language in use relies to a great extent on prefabricated word patterns (*how are you, it should be noted that*), and that the ability to use them appropriately is a significant component in successful language acquisition.

The advance of corpus methodologies, along with the rapid development of technology and concordancing software has enabled a deeper look into these patterns. Unlike previous research that was based on researchers' intuition in the identification of multi-word sequences (e.g. Nattinger and DeCarrico, 1992), in the last couple of decades the research has been empirical rather than intuitive, based on large corpora of electronically stored texts. Following Altenberg (1993, 1998) who used a frequency-driven, fixed word approach to identify multi-word sequences, Biber et al. identified *lexical bundles* as specific type of multi-word expressions, and defined them as "recurring sequence of three or four words" (1999, p. 990). Biber's lexical bundle approach has been applied in many subsequent corpus

Submitted February 2nd, 2019, accepted for publication April 21st, 2019

studies in EAP. These studies focus on exploring lexical bundles in university teaching and textbooks (Biber, Conrad, and Cortes, 2004), history and biology journals and students' writings (Cortes, 2004), conversation and academic prose (Conrad and Biber, 2005), university spoken and written registers (Biber and Barbieri, 2007) academic writing and disciplinary variation (Hyland, 2008), medical research articles (Jalali, 2015), academic lectures (Neely and Cortes, 2009), L1 and L2 academic writing (Chen and Baker, 2010). The findings from this research have demonstrated that the bundles are "important building blocks of discourse associated with basic communicative functions" (Biber, Conrad and Cortes, 2004, p. 400), that "different registers rely on different sets of lexical bundles" (Conrad and Biber, 2005, p. 69), and that grammatical structure of lexical bundles is a distinct characteristic of registers (Biber et al., 1999).

Given that bundles are discipline related, Conrad and Biber (2005) call for further study in more registers. To our knowledge, not much research has been done to study lexical bundles in the academic writing for economics. In addition, many of the studies focus on native English speakers, and more research is needed to gain insight into the use of bundles by non-native speakers. The present work attempts to fill in this existing gap by identifying the frequency and structure of 4-word lexical bundles used in economics research academic articles written by Macedonian scholars. In section 2 below an operational definition of lexical bundles is given, as well as brief overview of previous research regarding structure of bundles in the academic writing. Section 3 describes the methodology and corpus used for the present study. Section 4 analyses the structure of lexical bundles in economics research articles written by Macedonian speakers of English. The conclusion provides our final remarks as well as teaching implications regarding bundles in the register of economics.

2. Theoretical Background and Research Question

The term "lexical bundle" was first introduced in the extensive Longman Grammar of Written and Spoken English (Biber et al., 1999) to identify the recurrent multi-word sequences found in the Longman Spoken and Written English Corpus of conversation and academic prose. The definition states that lexical bundles are "combination of words that in fact recur most commonly in a given register" (Biber et al., 1999, p. 992). Lexical bundles have several specific characteristics that distinguish them from the other multi-word sequences, such as idioms and collocations. By definition, they are extremely common, not idiomatic in meaning, and not perceptually salient (Biber and Barbieri, 2007, p. 269). They are clusters of 3-6 contiguous words identified in electronic corpora exclusively on the basis of their frequency in a certain register, may be parts of clauses (*I don't want to*) or phrases (*in the case of*), and usually do not represent complete structural and semantic units. Biber et al. (1999) suggest that 15% of the bundles in conversation are structurally complete, while only 5% in the academic prose are complete structural units.

Biber, Conrad and Cortes (2004, p. 399) found that although "they are not the kinds of grammatical structures recognized by traditional linguistic theory, most lexical bundles do have well-defined structural correlates", and provided a structural classification of the bundles found in the Longman corpus, "taking into account the initial elements of the bundle and its overall structure" (Conrad and Biber, 2005, p. 60). Many scholars working in this field (e.g. Cortes, 2004; Hyland, 2008, Chen and Baker, 2010) have applied this

structural taxonomy to group the bundles in various registers, especially in the academic writing, and came to conclusion that the "principle structures of bundles differ across fields" (Hyland, 2008, p. 10). The bundles in the academic prose are "phrasal rather than clausal" (Biber, Conrad and Cortes, 2004, p. 382), with 70% of the bundles consisting of noun phrase expression (*the nature of the*), or a sequence that bridges across two prepositional phrases (*as a result of*). On the contrary, bundles in conversation are "parts of declarative clauses or questions" (Conrad and Biber, 2005, p. 63). Hyland (2008) states that most bundles in academic writing are parts of noun and prepositional phrases, and reports different patterns across disciplines: social sciences (business studies and applied linguistics) employ more bundles beginning with a prepositional phrase, while science and engineering texts use more passive bundles. Jalali (2015) found that the largest structural category of lexical bundles in medical research articles was prepositional phrases. Cortes (2004) also reports that the majority of the bundles used in academic history writing in English and Spanish are prepositional phrases. The current study follows this line of research, through a frequency-driven analysis of the structure of lexical bundles in the register of economics research articles written by non-native speakers of English. It focuses on the following research questions:

- What are the most frequent 4-word lexical bundles in the economics research articles written by Macedonian scholars, and what are their structural patterns?
- How do lexical bundles in economics research articles compare with the lexical bundles found by Hyland (2008) in research articles, PhD dissertations, Ma/MSc thesis from 4 disciplines?

3. Data and Methodology

In this research the lexical bundle method is used, that requires multi-word sequences to be identified in an electronic corpus, with priority given to frequency, fixedness and sequences longer than two words (Conrad and Biber, 2005). It follows Biber et al.'s (1999) definition and structural taxonomy of lexical bundles. The study incorporates register perspective, and it focuses on 4-word lexical bundles "because they are far more common than 5-word strings and offer a clearer range of structures and functions than 3-word bundles" (Hyland, 2008, p. 8). It adopts a corpus-driven approach, meaning that it is inductive. No multi-word sequences have been pre-selected (as in the case of corpus-based research), and "the linguistic constructs themselves emerge from analysis of a corpus" (Biber, 2009, p. 276).

3.1. Corpus used for the study

For the needs of the research an original corpus was compiled (ERAC, or economics research articles corpus), representative of the register of academic research articles in the field of economics. The articles were published in the period from June 2011 to June 2017, in the "Economic Development", and the "CEO Journal of Economics", both of which international journals of economics based in Skopje, Macedonia. The articles were downloaded in the period from January to March 2018, from the official websites of the journals. Only articles written by Macedonian scholars were included in the analysis. An effort was made to include approximate number of tokens from both journals. The "Economic Development" articles made up 60 files, 188,903 word tokens, and 8,128 word

types, while the articles from the "CEA Journal of Economics" made up 40 files, 189,299 word tokens and 8,356 word types. The total of 100 files was included in the corpus, containing 378,202 tokens.

The articles were published in a pdf format, and had to be converted in Word. Descriptive metadata for each file was created, containing the following information: title of the journal, date of issue (month, year), names of the authors, type of the article (e.g., original scientific paper), reference number and title of the article. After the conversion, all the articles were manually cleaned from the data irrelevant for linguistic analysis (names of authors, headings, tables, graphs, footnotes, references, page numbers and formulas were removed). Taking into consideration the fact that texts in the genre of economics would inevitably contain numerical data, the digits that were an integral part of the discourse (e.g., years, quantities, percentage) were not removed, in order not to damage the integrity of the texts. The next step included conversion of the Word documents into a plain text format using Unicode 8 encoding, in order to make the files readable by the concordancing software. The software used for the analysis is AntConc (Version 3.5.7) [Windows] 2018, which is a concordancing tool developed by Anthony Laurence. In the global settings of the software, the token definition was set only to *letter*, in order to avoid counting numbers and other symbols as tokens.

The function 'n-grams' of the software was used to identify 4-word lexical bundles in the corpus, using their frequency as a first identification criterion. The second identification criterion is the range, i.e. the distribution of lexical bundles in the corpus files, to avoid idiosyncratic use of language (Biber and Barbieri, 2007). In order to qualify as a lexical bundle, the word string must occur frequently in a specific register within a specific range, although Conrad and Biber (2005) point out that the frequency cut-off if arbitrary. Different cut-off frequencies and distribution criteria have been applied for identification of bundles in various studies, ranging from 10 bundles per million words in a register, spread across 5 different texts (Biber et al., 1999), 20 per million words across 10% of the texts (Hyland, 2008), 40 per million words in at least 5 different texts (Biber, Conrad and Cortes, 2004), 40 per million words in 20 different texts (Biber and Barbieri, 2007).

In this study we followed Hyland (2008), setting the frequency at 20 bundles per million words. For our corpus the converted raw-frequency threshold is 7.56, which was rounded up to 8 in order to be recognizable as a value by the concordancer. The distribution threshold was set at 10, i.e. the word-string must occur in 10% of the texts in order to be considered a bundle.

4. Analysis and Results

The software yielded 162 lexical bundles that met previously established criteria. In line with Chen and Baker (2010), lexical bundles containing proper nouns and context-based bundles (e.g. the ministry of finance, in the European Union) were neglected. After the exclusion of these bundles, a total of 141 different bundles were included in the analysis, and the obtained results were compared with Hyland's (2008) 4 sub-corpora. As shown in Table 1, economics articles contain fewer bundles compared with electrical engineering, and more bundles than biology, while the number of bundles in ERAC and social sciences is similar.

Table 1 Comparison of lexical bundles across disciplines in Hyland (2008) and ERAC

Corpus	Words in the corpus	Different bundles	Total cases of bundles	Words in bundles
ERAC	378,202	141	3583	3.8%
Electrical engineering	632,500	213	4562	3.5%
Business Studies	844,400	144	3728	2.2%
Biology	794,100	131	2909	1.7%
Applied Linguistics	1,129,400	141	4631	1.9%

Since our corpus has 378.202 tokens, we applied the normalisation formula (Biber and Barbieri, 2007), in order to make the bundles comparable with previous research. The most frequent 24 bundles have a normalized frequency higher than 100 per million words. The bundle *on the other hand* is the most frequent bundle in the corpus, with a raw frequency of 110, occurrence in 54 of the corpus files, and normalized rate of 290 times per million words. Table 2 shows the 20 most frequent bundles in ERAC.

Table 2 The 20 most frequent lexical bundles in ERAC

4-word bundles	Raw frequency	Normalized frequency	Range
on the other hand	110	290	54
as a result of	108	285	50
as well as the	88	232	44
a result of the	66	174	35
at the same time	62	163	34
the size of the	62	163	25
the total number of	62	163	25
on the basis of	61	158	26
in the process of	56	148	29
the development of the	56	148	23
in the case of	52	137	20
at the end of	49	129	21
is one of the	48	126	38
in accordance with the	44	116	24
in the banking sector	44	116	17
the analysis of the	44	116	31
the results of the	44	116	26
small and medium sized	42	111	12
the implementation of the	42	111	25
in terms of the	41	108	26

The comparison with Hyland (2008) revealed that total of 40 bundles from all bundles in ERAC are shared with Hyland's overall corpus. ERAC and social sciences corpora share 49 bundles, while ERAC and pure and applied sciences corpora have 31 bundles in common. Bundles that occur in all five disciplines are: *on the other hand, as well as the, at the same time, in the case of,* and *the results of the.* Table 3 summarizes this information.

Table 3 Shared bundles in Hyland (2008) and ERAC

Lexical bundles	ERAC	Biology	Electrical Engineering	Applied Linguistics	Business Studies
on the other hand	*	*	*	*	*
as a result of	*	*		*	*
as well as the	*	*	*	*	*
a result of the	*	*			
at the same time	*	*	*	*	*
the size of the	*		*		*
the total number of	*				*
on the basis of	*			*	*
in the process of	*			*	
in the case of	*	*	*	*	*
at the end of	*	*		*	*
is one of the	*	*	*	*	
the results of the	*	*	*	*	*
in terms of the	*		*	*	*
it is necessary to	*		*		
one of the most	*			*	*
the end of the	*	*		*	
the structure of the	*		*		
the beginning of the	*	*		*	
the fact that the	*	*		*	*
in the form of	*		*	*	*
the other hand the	*		*	*	*
the role of the	*			*	
to the fact that	*			*	
it is important to	*			*	*
for the purpose of	*			*	
in relation to the	*			*	
is due to the	*	*	*		
that there is a	*			*	
in the number of	*				*
is based on the	*		*		
the quality of the	*	*			*
at the beginning of	*	*		*	
in the context of	*			*	*
the performance of the	*		*		
the purpose of the	*			*	
a wide range of	*	*		*	*
the extent to which	*				*
it can be seen	*		*		
on the one hand	*			*	

The comparison also revealed that 12 of the 20 most frequent bundles in ERAC are found among the 20 most frequent bundles in Hyland's overall corpus. If compared with the sub-corpora, ERAC and social sciences have 17 bundles in common, while ERAC and applied and pure sciences share 11 bundles. Table 4 presents the bundles shared in all 5 disciplines, with the shared bundles given in bold.

Table 4 Shared bundles among the 20 most frequent in Hyland (2008) and ERAC

ERAC	**on the other hand**, **as a result of**, **as well as the**, a result of the, **at the same time**, **the size of the**, the total number of, **on the basis of**, **in the process of**, the development of the, **in the case of**, **at the end of**, **is one of the**, in accordance with the, in the banking sector, the analysis of the, **the results of the**, small and medium sized, the implementation of the, **in terms of the**
Biology	in the presence of, in the present study, **on the other hand**, the end of the, **is one of the**, **at the end of**, it was found that, at the beginning of, **as well as the**, **as a result of**, it is possible that, are shown in figure, was found to be, be due to the, **in the case of**, is shown in figure, the beginning of the, the nature of the, the fact that the, may be due to
Electrical Engineering	**on the other hand**, as shown in figure, **in the case of**, is shown in figure, can be seen that, can be used to, the performance of the, as a function of, is based on the, with respect to the, is given by equation, the effect of the, the magnitude of the, **at the same time**, in this case the, it is found that, **the size of the,** be seen that the, the accuracy of the, **as well as the**
Applied linguistics	**on the other hand**, **at the same time**, **in terms of the**, **on the basis of**, in relation to the, **in the case of**, in the present study, the end of the, the nature of the, in the form of, **as well as the**, **at the end of**, the fact that the, in the context of, **is one of the**, **in the process of**, **the results of the**, in terms of their, to the fact that, in the sense that
Business studies	**on the other hand**, **in the case of**, **at the same time**, **at the end of**, **on the basis of**, **as well as the**, the extent to which, the end of the, significantly different from zero, are more likely to, the relationship between the, the results of the, the hang seng index, the other hand the, in the context of, **as a result of**, the performance of the, hong kong stock market, is positively related to, are significantly different from

In the next step of the analysis, Biber et al.'s (1999) classification was used to group the bundles structurally, i.e. by grammatical parts. The bundles from ERAC fitted into the categories of bundles in the academic prose as suggested by Biber. The structure of the lexical bundles in ERAC is given in Table 5.

In line with what Biber et al. (1999) and Cortes (2004) reported in their research, the lexical bundles found in the ERAC are not grammatically complete units. The main category is 'noun phrases with of-phrase fragment', with 29.08% of all bundles found in the corpus. The bundles *a result of the, the size of the, the total number of* are the most frequent in this category. The majority of the bundles follow the structure 'definite article + noun + preposition *of* + definite article' (*the development of the*). This category is the most common in Hyland's overall corpus as well, with the largest percentage in the sub-corpus of business studies (28.5%). The sub-corpus of electrical engineering employs the least percentage of bundles from this category, i.e. 22.3%. 'Other prepositional phrase fragment' (*at the same time*) is the second most frequent category, containing 17.74% of all bundles. Compared across corpora, the use of these bundles is similar in the sub-corpus of business studies (19.7%). This structure is the most common in the field of applied linguistics (24.4%), while biology and electrical engineering use only 13.7% and 11.6% of these bundles, respectively. The number of 'prepositional phrases with embedded of-phrase fragment' (*as a result of*) comes third in ERAC, represented by 15.61%. The prepositions used in these two categories are *in, of, for, with, at, as, to, on,* with the prepositions *in* and *of* being the most common

Table 5 Structure of 4-word bundles in ERAC with examples

Types of bundles in %		
PHRASAL 70.95%	NP with *of*-phrase fragment (*the analysis of the*)	29.08
	Other PP fragment (*in accordance with the*)	17.74
	PP with embedded *of*-phrase fragment (*in the process of*)	15.61
	NP with other post-modifier fragment	8.52
	1. NP with a post nominal clause fragment (*the fact that the*)	
	2. NP with a prepositional phrase fragment (*the changes in the*)	
Other expressions (*as well as in*)		7.80
CLAUSAL 21.25%	VP + *that*-clause fragment	4.96
	1. VP + *that*-clause (*can be concluded that*)	
	2. *that*-clause (*that there is a*)	
	Adverbial clause fragment (*in order to improve*)	4.96
	Anticipatory *it* +	4.25
	1. verb phrase (*it is expected that*)	
	2. adjective phrase (*it is important to*)	
	Passive verb + PP fragment (*be taken into account*)	2.84
	Copula *be* + noun phrase (*is the fact that)*	2.84
	Verb/adjective + *to*-clause fragment (*to be able to*)	0.70
	Pronoun/noun phrase + *be* + (*this paper is to*)	0.70

introductory prepositions. The percentage for this category in ERAC is the closest with the one in business studies (16.0%), while the greatest difference was found with the sub-corpus of electrical engineering, containing only 7.9%. The category 'noun phrases with other post-modifier fragment' is represented by 8.52%, and it contains two sub-groups: 'noun phrase with a post nominal clause fragment', covering bundles such as *the extent to which*, and 'noun phrase with a prepositional phrase fragment', with bundles like *the results from the*. The percentage of 'noun phrase with other post-modifier fragment' in the economics articles corpus is slightly lower compared to the other four disciplines in Hyland's corpus. Nonetheless, if taken together with the 'noun phrases with of-phrase fragment' category, the total percentage of lexical bundles containing noun phrases is 37.6%, representing the majority of all bundles in the corpus. The group 'other expressions' encompasses bundles that do not belong to any of the other categories, and is represented with 7.80%. It contains adverbial phrase bundles like *as well as the*, adjectival phrase bundles, such as *small and medium sized*, conjunction + prepositional phrase bundles, such as *and at the same*, and others.

Table 6 Structure of lexical bundles in Hyland (2008) and ERAC (%)

Types of bundles	ERAC	Biology	Electrical engineering	Applied linguistics	Business studies
NP + *of*	29.08	23.7	22.3	22.9	28.5
NP + other modification	8.52	9.4	10.8	9.6	12.4
PP + *of*	15.61	9.2	7.9	19.9	16.0
Other PP	17.74	13.7	11.6	24.4	19.7
Anticipatory *it*	4.25	6.3	8.4	5.6	4.5
Passive verb + PP	2.84	31.3	29.8	6.9	9.0
Other expressions	7.80	6.4	9.2	10.7	9.9

Only 22.7% of all bundles in ERAC include a verb, which supports Conrad and Biber's (2005) claim that verbs are not commonly found in academic prose. There are 14 lexical verbs contained in the bundles, and the bundle *is one of the* is the only one among the 20 most frequent that contains a verb. The most common bundles among the clausal bundles are from the categories 'verb phrase + *that*-clause fragment' and 'adverbial clause fragment', each taking up 4.96% of the bundles in ERAC. The 'verb phrase + *that* clause fragment' contains two sub-categories: 'verb phrase + *that*-clause', with bundles such as *having in mind that*, and '*that*-clause', covering bundles like *that there is no*. All the bundles in the group 'adverbial clause fragment' incorporate *to*-clause (*in order to achieve*). The 'anticipatory *it* +' takes up 4.25% of the bundles, represented in structures that incorporate a verb phrase (*it can be concluded*), and an adjective phrase (*it is necessary to*). The comparison for this structure showed that the percentage in ERAC is the closest with the one in business studies (4.5%), while the biggest discrepancy was found with the corpus of electrical engineering (8.4%). The category 'passive verb + prepositional phrase fragment' (*is based on the*) takes up 2.84% of the bundles, which is the lowest in comparison with the other disciplines. This category reflects the biggest difference between social sciences (9.0% in business studies and 6.9% in applied linguistics) and pure and applied sciences, which tend to use much more passive bundles (31.3% in biology and 29.8% in electrical engineering). The low percentage of these bundles in ERAC goes in line with Hyland's (2008) finding that social studies do not tend to use passive bundles, unlike the applied and pure sciences which employ significantly more passive bundles. Byrd and Coxhead (2010) also report lack of passive bundles in their study of academic writing and teaching. 'Copula *be* + noun phrase/adjective phrase' is represented by 2.84% as well, containing bundles that incorporate only noun phrases (*is a result of*). The least frequent bundles in ERAC are from the categories 'verb/adjective + to-clause fragment' (*to be able to*) and 'pronoun/noun phrase + be +' (*this paper is to*), represented by 0.70% each, or just one occurrence in the corpus. The complete structural classification of the bundles in ERAC and Hyland's corpus is presented in Table 6.

5. CONCLUSION

The findings of the present study demonstrate that Macedonian scholars in the discipline of economics rely on 4-word lexical bundles. Economics research articles show similarity in the use of bundles with applied linguistics and business studies, which is not surprising having in mind that they all belong to the social sciences. Several bundles among the 20 most frequent (*the development of the, in accordance with the, in the banking sector, the analysis of the, small and medium sized* and *the implementation of the*) are not found in the other corpora, meaning that they are specific for the register of economics academic prose, which lends support to the concept that bundles are register specific (Biber et al., 1999). These bundles could be taken into consideration by curriculum developers in EAP, especially in developing teaching materials for English for Economics. In contrast, the bundles: *on the other hand, as well as the, at the same time, in the case of,* and *the results of the* occur among the 20 most common bundles in all five disciplines, and can potentially be included in general EAP courses.

Regarding the structure, the majority of the bundles used in this academic register are phrasal, showing clear dominance over clausal bundles. The phrasal nature of the

academic prose has been demonstrated in several studies (Biber et al. 1999, Conrad and Biber, 2005, Hyland, 2008, Byrd and Coxhead, 2010). The results of the current research support this finding, showing that 70.95% of the bundles in the corpus of economics research articles are noun phrases and prepositional phrases. In the comparison across corpora, the biggest difference was found with the biology and electrical engineering corpora in the use of passive verb bundles.

From a pedagogical point of view, the results of this research suggest that noun and prepositional phrases should be given more significance in course books. Traditionally, the verb phrase has been given primary importance in the syllabuses, but the corpus based research has discovered the omnipresence of the phrasal structures in various disciplines, and their supremacy over clausal structures. The teaching of phrasal bundles needs to be emphasized in English for Economics as a part of EAP, and should move from teaching simple noun and prepositional phrases towards more complex ones, found in a corpus containing the authentic language of the specific register, i.e. target discourse.

One of the possible limitations of this study could be the size of the corpus, since smaller corpora might produce larger number of bundles compared with larger corpora from the same register (Hyland, 2012). Another factor that might have influenced the results is the impact of the native language over the use of bundles, which was not taken into consideration in this research. Future research should move in the direction of studying the possible methods and techniques to introduce and teach the structure of lexical bundles in EAP classrooms.

REFERENCES

Altenberg, Bengt. 1998. “On the Phraseology of spoken English: The evidence of recurrent word-combinations.” In *Phraseology: Theory and Applications*, edited by Anthony Paul Cowie, 101-122. Oxford: Oxford University Press.

Altenberg, Bengt. 1993. “Recurrent word combinations in spoken English.” In *Proceedings of the fifth Nordic association for English studies conference*, edited by Julian D'Arcy, 17-27. Reykjavik: University of Iceland.

Anthony, Laurence. 2018. AntConc (Version 3.5.7) [Windows]. Tokio, Japan: Waseda University. Available from http://www.laurenceanthony.net/software.

Biber, Douglas. 2009. “A corpus driven approach to formulaic language.” *International Journal of Corpus Linguistics*, 14:3, 275-311. doi 10.1075/ijcl.14.3.08bib issn 1384–6655 / e-issn 1569–9811.

Biber, Douglas, and Federica Barbieri. 2007. “Lexical bundles in university spoken and written registers.” *English for Specific Purposes*, 26(3), 263–286.

Biber, Douglas, Susan Conrad, and Viviana Cortes. 2004. “If you look at...: Lexical bundles in university teaching and textbooks.” *Applied Linguistics, 25L*, 371–405.

Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad, and Edward Finegan. 1999. *Longman grammar of spoken and written English*. London, England: Longman.

Byrd, Pat, and Averil Coxhead. 2010. “On the other hand: Lexical bundles in academic writing and in the teaching of EAP.” *University of Sydney Papers in TESOL*, 5, 31–64.

Chen, Yu-Hua, and Paul Baker. 2010. “Lexical bundles in L1 and L2 academic writing." *Language Learning and Technology*, 14(2), 30–49.

Conrad, Susan, and Douglas Biber. 2005. "The frequency and use of lexical bundles in conversation and academic prose." *Applied Linguistics Faculty Publications and Presentations*.
https://pdxscholar.library.pdx.edu/cgi/viewcontent.cgi?article=1002&context=ling_fac

Cortes, Viviana. 2004. "Lexical bundles in published and student disciplinary writing: Examples from history and biology." *English for Specific Purposes*, 23(4), 397–423.

Firth, John Rupert (1951). *Modes of meaning.* Essays and Studies (The English Association). 118-149.

Hyland, Ken. 2012. "Bundles in academic discourse." *Annual Review of Applied Linguistics*, *Vol. 32*, 150–169.

Hyland, Ken. 2008. "As can be seen: Lexical bundles and disciplinary variation." *English for Specific Purposes, 27*, 4–21.

Jalali, Zahra Sadat, Mohammad Raouf Moini, and Mohamad Alaee Arani. 2015. "Structural and Functional Analysis of Lexical Bundles in Medical Research Articles: A Corpus-Based Study." *International Journal of Information Science and Management Vol. 13*, No. 1, 51-69.

Lewis, Michael. 1993. *The lexical approach*. Language teaching publications.

Nattinger, James R., and Jeanette S. DeCarrico. 1992. *Lexical phrases and language teaching*. Oxford: Oxford University Press.

Neely, Elizabeth, and Cortes Viviana. 2009. "A little bit about: analyzing and teaching lexical bundles in academic lectures." *Language Value. Vol. 1*, N° 1, 17-38.

Schmitt, Norbert. 2004. *Formulaic sequences*. Amsterdam, Netherlands: John Benjamins.

Sinclair, John. 1991. *Corpus, concordance, collocation*. Oxford: OUP.