The structural features of English lexical bundles in academic essays written by Filipino college students

Rey John Castro Villanueva

Corresponding author
Rey John Castro Villanueva
Mariano Marcos State University - Batac, Ilocos Norte, Philippines
arjaycastrovillanueva25@gmail.com

Abstract

Lexical bundles are part of any writing production. As these recurrent word combinations appear in various written outputs, many linguists, students, and researchers have become interested in exploring such linguistic expressions. However, most of them who examined the features of lexical bundles used computer corpora collected from native speakers of English. Therefore, little is known about lexical bundles produced by nonnative users of English, particularly the Filipinos who use English as their second language in government, business, religion, and education domains. Using the structural categories developed by Biber et al. (1999), this study examined the grammatical characteristics of three- and four-word English lexical bundles in a 100,000-word corpus produced by college students from different Philippine colleges and universities. The results revealed that the students seemed to have insufficient knowledge of these language units because a majority of them used obscure and unsophisticated bundles. Thus, the findings of the study imply that the students should familiarize themselves with these English lexical bundles, for these word combinations serve as basic elements or building blocks for their written outputs.

Keywords: Lexical bundles, structural taxonomy, academic essays, Filipino college students

1.   Introduction

Lexical bundle is an aspect of an academic register that was first described in Biber, Johansson, Leech, Conrad, and Finegan's (1999) *Longman Grammar of Spoken and Written English*, a monumental scholarly work completely based on the 100-million-word *British National Corpus*. Biber and his colleagues defined lexical bundles as sequences of frequently co-occurring words in particular registers. They are present in written registers, and they are considered "basic building blocks for constructing […] written discourse" (Biber & Conrad, 1999, p. 188). Moreover, they stated that "both conversation and academic prose use a large stock of different lexical bundles" (p. 993). Afterward, such a claim became a cornerstone for further studies on lexical bundles in university registers.

Using the compositions of freshman L1 speakers of English, Cortes (2002) made her own analysis on English lexical bundles. Her corpus, consisting of 360,704 running words, revealed a total of 93 different bundles that were structurally and functionally examined. While the structural scrutiny showed that such language units were similar to those used in other academic prose, the functional analysis, on the other hand, explained that these bundles served as temporal or locative markers, which created redundancy in the writings of the students.

Two years later, Cortes (2004) conducted another study that compared the written outputs (904,376 words) of university students who are native speakers of English with published journal articles (1,992,531 words). Her corpus contained two main disciplines – history and biology. The research findings revealed that the students seldom use lexical bundles identified in the second set of corpus, i.e., published journal articles.

In the same year, Biber, Conrad, and Cortes (2004) conducted another comprehensive research using the TOEFL 2000 Spoken and Written Academic Language (T2K-SWAL) Corpus. They found that the lexical bundles in the said corpora differed dramatically from other linguistic features and that university lectures used twice as many lexical bundles than conversation and four times as many lexical bundles as textbooks.

Nesi and Basturkmen (2006), on the other hand, examined the cohesive role that lexical bundles take part in the TOEFL 2000 Spoken and Written Academic Language (T2K-SWAL) Corpus. The study found that lexical bundles can play a discourse-signaling role in lectures. Moreover, Nesi and Basturkmen noted that it is crucial to raise students' awareness of such a function of lexical bundles.

Kim (2009) also became interested in investigating lexical bundles; however, she explored a language other than English. She examined a large corpus of Korean texts, which was consisted of academic and conversation prose. Her findings revealed that lexical bundles are significant expressions in Korean texts because they act as discourse frames for new information.

In their attempt to investigate the influence of nonnativeness, Chen and Baker (2010) analyzed the published academic texts (FLOB-J) and two corpora of student academic writing (BAWE-EN and BAWE-CH). Their study revealed that while FLOB-J displayed the widest range of lexical bundles, the L2 BAWE-CH, on the other hand, showed the smallest.

Finally, Salazar (2010) conducted a study, which she delineated as a frequency-driven investigation of the occurrence as well as the grammatical and functional taxonomies of lexical bundles with verbs in Philippine and British scientific English. The findings of this quantitative analysis showed a lower amount of verbal lexical bundles in the nonnative (Philippine) computer corpus compared with the native (British) corpus. Such a research, therefore, provided a description of the structural and functional similarities and differences between the British and Filipino corpora of medical articles with regard to the use of lexical bundles.

The studies presented above show that up to the present, lexical bundles have not been explored in the Philippine context, taking into consideration written corpus compiled among Filipino university students. The present research, therefore, is an attempt to fill that

gap. The researcher believes that by examining English lexical bundles used in ESL writing classrooms, the students may be given opportunities to enhance their writing skills and become more competent users of English. As Rafiee and Keihaniyan (2013) opined, language learners need to know the structures of lexical bundles to be able to improve their writing skills.

It should be made clear that there are other types (e.g., two-, five-, and six-word) of English lexical bundles; however, this research preferred to focus on three- and four-word lexical bundles only because many scholars claim that these are the most researched bundle types and that such have been used in several related studies. Moreover, Salazar (2011) expressed that the longer the bundle, the lower is its frequency. Cortes (2004), on the other hand, argued that these types of lexical bundles provide researchers with more obvious varieties of structures to analyze.

The main purpose of this study was to investigate the structural features of English lexical bundles in academic essays written by Filipino college students. Specifically, it sought answers to the following questions:

1.   What are the three- and four-word English lexical bundles found in academic essays written by the students?
2.   What are the structural types of these lexical bundles found in the study?

2.   Method

2.1   Study Corpus

The corpus was built using the academic essays written by 130 fourth-year college students enrolled in two writing courses, i.e., Business English and Technical English. The students were tasked to write academic essays about any social issue in the Philippines during their vacant periods in order to avoid class interruptions. Moreover, senior students were chosen as informants of this study because it was assumed that they have already attained the intermediate or the [nearly] advanced level of English proficiency. Thus, they may be expected to be models and provide models of what is acceptable and prestigious English in Philippine colleges and universities.

The corpus constitutes 130 essays with a total of 100,000 running words. In a study, Biber (2006) stated that a corpus must be large enough to adequately represent the occurrence of certain features under study; however, Bowker and Pearson (2002) and Pierini (2009) once argued that small corpora could be reliable and representative, especially when dealing with domain-specific languages.

_____

2.2   Operationalization

In large written corpora, the normalized frequency threshold generally ranges from 20 to 40 per million words. For the present study, the researcher took a sort of conservative approach by setting up the frequency point at occurring 20 times per hundred thousand words because of the small size of the corpus. Furthermore, in order to avoid the students' idiosyncratic use, a lexical bundle (both three- and four-word) must be used in at least three different academic essays. In terms of retrieving lexical bundles, the Antconc 3.2.4w (Anthony, 2007), a linguistic computer software, was utilized, and the following steps were strictly followed in generating English lexical bundles from the corpus:

a.   All the academic essays were digitized, and each of them was saved in Plain Text format as the Antconc 3.2.4w cannot process a corpus saved in other document formats.

b.   After the digitization process, the 'N-Grams' control of the Antconc 3.2.4w was used to generate both three- and four-word English lexical bundles from the corpus.

After the operation of AntConc 3.2.4w based on the abovementioned settings, a list of three- or four-word English lexical bundles were culled, and each lexical bundle in the list was manually checked to ascertain whether or not it appeared in more than three different texts in the corpus. A recurrent word combination that occurred in less than three texts was not considered as a lexical bundle and was, therefore, removed.

2.3   Structural Categories of Lexical Bundles

The three- and four-word lexical bundles were classified according to their structures or grammatical types using the taxonomy shown in Table 1 below. This structural classification of lexical bundles in *Longman Grammar of Spoken and Written English* was used in this research because as what Beng and Keong (2014) affirmed, such a "structural classification has been widely relied on in the studies of lexical bundles" (p. 128).

Table 1
Structural types of lexical bundles

Categories		Subcategories	Sample Bundles
NP-based	1)	noun phrase with *of*–phrase fragment	*the end of the, the form of a*
	2)	noun phrase with other post-modifier fragment	*the way in which, the degree to which*
	3)	pronoun/noun phrase + *be* (+...)	*this is not the, there are a number*
PP-based	4)	prepositional phrase with embedded *of* – phrase fragment	*in the context of, at the end of the*
	5)	other prepositional phrase fragment	*as in the case, to the fact that*
VP-based	6)	anticipatory *it* + verb phrase/ adjective phrase	*it can be seen, it is possible to*
	7)	passive verb + prepositional phrase fragment	*is based on the, is shown in figure*
	8)	copula *be* + noun phrase/ adjective phrase	*is one of the, is due to the*
	9)	(verb phrase +) *that* – clause fragment	*should be noted that, that it is not*
	10)	verb/adjective phrase +) *to* – clause fragment	*to say that the, is likely to be*
	11)	adverbial clause fragment	*as we have seen, as we shall see*
Others	12)	Other expressions	*as well as the, the presence or absence*

3.   Results and Discussion

3.1   An Inventory of Three- and Four-word English Lexical Bundles in Students' Academic Essays

Using the concordancing software, Antconc 3.2.4w, the following lexical bundles (See Table 2.) were generated. As shown by the figures, a total of 31 lexical bundles were identified. A majority of them are three-word lexical bundles since Antconc 3.2.4w found only one four-word lexical bundle. This four-word lexical bundle, *is one of the*, which appeared in 13 texts, was employed 33 times by the students.

Table 2
A detailed list of the three- and four-word English lexical bundles

Three-word	Frequency	Four-word	Frequency
one of the	80	is one of the	33
a lot of	78		
because of the	49		
the fact that	42		
of the country	39		
of the Philippines	39		
that it is	37		
the Philippine is	37		
to have a	36		
is one of	35		
because it is	34		
be able to	32		
in the country	31		
it is a	30		
the use of	29		
of the people	28		
that they are	26		
it is not	25		
should not be	25		
in order to	23		
that there are	23		
according to the	22		
in the world	22		
it comes to	22		
they do not	22		
due to the	21		
of the most	21		
there is a	21		
in our country	20		
the people who	20		

On the other hand, the three-word lexical bundles that occurred more than 38 times in the corpus include: *one of the, a lot of, because of the, the fact that,* and *of the country.*

The three-word lexical bundle, *one of the*, was used 80 times by the informants and appeared in 25 texts. Such a lexical bundle is followed by *a lot of*, which occurred 78 times, and was used in nine (9) texts. *Because of the* ranked third with 49, and it appeared in 20 texts. While, *the fact that* and *of the country* are the fourth and fifth most frequently used lexical bundles, appearing 42 and 39 times in 20 and 14 texts, respectively.

3.2   Structural Analysis of Three- and Four-word Lexical Bundles in Academic Essays

Table 3 presents the quantitative analysis of the structural taxonomies of three- and four-word English lexical bundles present in the corpus. It should be noted that most of the recurrent word combinations, as what Biber et al. (1999) and Cortes (2004) argued, are not grammatically complete units. However, unlike two-word lexical bundles, three- and four-word lexical bundles are easier to categorize structurally because they provide a concise layout of grammatical structure.

The identified three- and four-word lexical bundles were, therefore, examined in succession, and then each of them was matched with the structural categories developed by Biber et al. (1999). Table 3 reveals the results of the structural classification procedure. All the structural categories, which appeared in the match-up operation, are all exhibited in the said table with their corresponding lexical bundles. Furthermore, for a better presentation of the said findings, the percentage and the rank of each of the taxonomies are also stated.

On the other hand, it should be made clear that no instance of 'prepositional phrase with embedded, *of* – phrase fragment,' 'passive verb + prepositional phrase fragment,' and 'adverbial clause fragment' was found in the structural analysis of the bundles. Meanwhile, the top two structural types that seem to be commonly employed by the students whenever they write academic essays are 'other prepositional phrase (fragment),' which covered 29% of the entire lexical bundles generated by the computer-operated text-analysis tool, followed by 'other expressions,' which has a percentage (25%) that is slightly different from the proportion earned by the previous category. The two other categories that appeared many times in the match-up procedure are 'noun phrase with of-phrase fragment' and 'anticipatory it + verb phrase/adjective phrase,' which ranked third (10%) and fourth (9.6%), respectively. These results demonstrate that the English lexical bundles in the academic corpus are phrasal rather than clausal in agreement with the findings of previous studies (e.g., Biber et al., 1999) that academic writing, unlike some registers such as conversation and teaching, are characterized as including more phrasal rather than clausal English lexical bundles.

Furthermore, it is worth noting that three of the nine structural taxonomies all ranked fifth (6.4%) as most of the subjects chosen for this inquiry marginally preferred to use lexical bundles that fall under these categories: 'copula *be* + noun phrase/adjective phrase,' '(verb phrase +) *that* – clause fragment,' and '(adjective/verb +) *to* – clause fragment.'

Table 3
Structural classifications of lexical bundles

Taxonomy	Lexical Bundle	Percentage	Rank
Noun phrase with *of*-phrase fragement	*one of the, a lot of, the use of*	10	3
Noun phrase with other post-modifier fragement	*the fact that*	3.2	8.5
Copula *be* + noun phrase/adjective phrase	*is one of, is one of the*	6.4	5.5
(Verb phrase +) *that* – clause fragment	*that they are, that it is*	6.4	5.5
(Adjective/verb +) *to* – clause fragement	*be able to, to have a*	6.4	5.5
Pronoun/noun phrase + *be* (+...)	*there is a*	3.2	8.5
Other prepositional phrase (fragment)	*of the country, of the Philippines, in the country, of the people, in the world, of the most, in our country*	29	1
Anticipatory *it* + verb phrase/ adjective phrase	*it is a, it is not, it comes to*	9.6	4
Other expressions	*because of the, the Philippines is, because it is, should not be, according to the, they do not, due to the, the people who*	25	2

A more detailed look at Table 3 indicates that 25% of the lexical bundles used by the informants in developing their essay topics are classified as 'other prepositional phrase (fragment).' These are *of the country, of the Philippines, in the country, of the people, in the world, of the most,* and *in our country*. Biber et al. (1999) claimed that these lexical bundles begin "with a prepositional phrase without an embedded *of* – phrase" (p. 1018), as in the following sentences:

[1]   These destructions can add up to poverty *of the country* because it can destroy different structures and plantation area and it can kill people.

[2]   This is because the government *of the Philippines* will have to solve these first for them to solve the major problem which is poverty.

_____

[3]   The primarily purpose of the bill is to solve the rise of population *in the country* named after the Pearl of the Orient.

[4]   Even though it is very expensive to buy, to maintain and to power a car, it is still preferable by most *of the people* because of the comfort it provides to the people.

[5]   Almost all the countries *in the world* prioritize education for their countrymen as they value the essence of having a better education for a better future.

[6]   It is one *of the most* controversial issues that can even lead to the fall of an entire country.

[7]   The people *in our country* specifically those in the lower level are most of the time neglected or taking for granted by the government.

Based on the structure of the sentences presented above, one may confirm that a majority of them refer to only one or two central topics. As the academic essays written by the students revolved around *social issues* in the Philippines, it is apparent then that the prepositions used instantaneously relate to 'the country' (the Philippines) or to 'the Filipinos.' Hence, such a finding substantiates the idea posited by Hyland (2008) that "most [lexical] bundles in academic writing are parts of [...] prepositional phrases" (p. 9).

Next to 'other prepositional phrase (fragment)' is 'other expressions.' The lexical bundles classified under this taxonomy occurred several times in the academic corpus. These lexical bundles have unclear grammatical attributes, which make the former unsuitable for any of the structural categories developed by Biber et al. (1999). The following sentences contain such bundles:

[8]   I know for a fact that we are not poor but we choose to be poor *because of the* way we are thinking nowadays have greatly affected our mindset towards being successful as a nation and as a country.

[9]   Although *the Philippines is* rich in resources, poverty cannot be eradicated because the country is prone to calamity, the Filipinos are idle, and the government is doomed to graft and corruption.

[10] The Filipino subject *should not be* abolished in college for the reason that the number of unemployed teachers and professors will increase.

[11] Still *according to the* same article, attention was given to the issue of the black market.

_____

[12] These problems will not get solved if *they do not* focus on this first.

[13] Many bills are still in discussion *due to the* Church's disapproval

[14] This involves emotions of *the people who* are hurt and *the people who* are not contended.

On the other hand, each of the italicized lexical bundles used in the three sentences below are classified as 'noun phrase with *of* – phrase.' Biber et al. (1999) noted that the lexical bundles, which normally fall under this category, provide a broad range of meanings. Moreover, they further explicate their point by positing that recurrent word combinations that belong to such taxonomy "are used for physical description, including identification of place, size, and amount" (p. 1015).

[15] Giving birth is *one of the* hardest parts yet wonderful thing here on Earth.

[16] They can do like *a lot of* job fairs and invest in *a lot of* businesses to solve unemployment.

[17] Another is that, shops sell and provide reusable eco-friendly bags to be able to reduce *the use of* plastic.

As can be seen in Table 3, 9.6% of the entire lexical bundles identified by Antconc 3.2.4w are categorized as 'anticipatory *it* + verb phrase/adjective phrase.' These lexical bundles placed under such a grammatical taxonomy are described by Biber et al. (1999) as those "that initiate extraposed structures" (p. 1019). In addition, Biber and his colleagues explained that these lexical bundles have two subdivisions: those controlled by an adjective phrase, and a fewer number controlled by a verb phrase. Here are sample extracts that show these lexical bundles.

[18] Cyberbullying can be considered as a misuse of people's right to the freedom of speech, and *it is a* way of degrading our peers.

[19] Although same sex marriage is viewed as unacceptable by the church, it should be legalized because *it is not* wrong to marry the same sex and *it is not* recognized nor prohibited by the Philippine law.

[20] This means that the employee would have a lot of advantage when *it comes to* using his/her healthcare.

_____

The informants who added English lexical bundles categorized as 'anticipatory *it* + verb phrase/adjective phrase' have the objective to convey 'possibility/likelihood,' 'importance,' and 'necessity' in their academic essays.

Following the top four grammatical types of recurrent word combinations involved in this research are the fifth ranked structures (i.e., 'copula *be* + noun phrase/adjective phrase,' '[verb phrase +] *that* – clause fragment,' and '[adjective/verb +] *to* – clause fragment'), which earned 6.4% of all the lexical bundles retrieved from the corpus.

For the first structure, 'copula *be* + noun phrase/adjective phrase,' Biber et al. (1999) asserted, "The lexical bundles in this category all begin with the copula *be* (or *may be*). There are two main sub-groups here, depending on whether the subject predicative is a noun phrase or adjective phrase" (p. 1021).

The sentences below consist of recurrent word combinations classified as 'copula *be* + noun phrase/adjective phrase.' Under this classification, the lexical bundles fall out as subject predicative to the copula *be* (Biber et al., 1999).

[21] Language is *one of the* important fundamentals that helps strengthen nationalism.

[22] Additionally, it is *one of the* most commonly used forms of bullying and is prevalent in the present times.

Another structural classification that ranked fifth is '(verb phrase +) *that* – clause fragment.' As implied by the name of the structure, the lexical bundles under such a grammatical type consist of the main clause verb and are tailed with *that* – clause, as in the following sentences:

[23] In fact, they are so devoted *that they* are willing to spend their country's entire budget on the military itself.

[24] Although it is not obvious *that it is* gravely important, some do not have an opportunity to be educated due to lack of money and government's assistance.

In some cases, the academic corpus also consists of recurrent word combinations that integrate *to* – clauses. Such lexical bundles are, as might be expected, categorized as '(verb/adjective +) *to* – clause fragment.' The succeeding strings of words were developed using this type of lexical bundles.

[25] One of the goals of the executive order 009 or the anti-plastic law is to *be able to* reduce the non-biodegradable materials and solid wastes that add up to the pollution of the city and clog the waterways which result to flooding in some areas.

_____

[26] Education is an absolute must and everyone has a right to in order *to have a* bright future.

Furthermore, Biber et al. (1999) affirmed that the lexical bundles of the above sentences have three main types. These are: (a) predicative adjective + *to* – clause; (b) (passive) verb phrase + *to* – clause; and (c) simple *to* – clause.

Based on the discussion above, one may contend that a majority of the student-informants enrolled at Philippine tertiary institutions and are considered to have already acquired an intermediate or [nearly] advanced level of English proficiency are not very much knowledgeable about English lexical bundles. Nevertheless, the results of the present research, albeit it applied a small corpus, seem to corroborate the findings of previous studies, particularly Salazar's (2010) research on lexical bundles, as 26% of the lexical bundles identified by the Antconc 3.2.4w were also present in the computer corpora produced by native speakers of English. This scenario, however, does not mean that the college students already know one-half of what the native speakers know about English lexical bundles. It should be noted that next to the top structural category, 'other prepositional phrase,' is the 'other expressions.' This means that a considerable quantity of the lexical bundles generated from the academic corpus are completely obscure, for they did not correspond to any of the grammatical typologies organized by Biber et al. (1999). Thence, such inability to use lexical bundles successfully in ESL writing classrooms may indicate that the informants involved in this study are not yet sharply exact in bringing forth unconstrained writing outputs.

4.   Conclusion

This study attempted to investigate three- and four-word English lexical bundles in academic essays written by Filipino senior college students from different Philippine colleges and universities. While there are other forms (e.g., two-, five-, and six-) of lexical bundles, this study focused only on three- and four-word lexical bundles, for these are the most researched types and have been used in several related studies. Apart from the identification of the most frequently used three- and four-word lexical bundles, the structural classification of each of the recurrent word combinations was also determined using the taxonomy developed by Biber et al. (1999).

Using the computer-operated text-analysis tool, Antconc 3.2.4w, a total of 31 recurrent word combinations were retrieved from a 100,000-word academic corpus. The size of the corpus used in this research was somewhat small; however, Fuster-Marquez (2014) claimed that "size is not necessarily the most relevant criterion in corpus building" (p. 91). A majority of the lexical bundles generated by the concordancing software are three-word lexical bundles, as the text-analysis tool identified only one four-word lexical bundle, i.e., *is one of the*, from the corpus. It is worth noting that 26% of the recurrent word combinations used by the informants in building up the corpus were also identified by previous studies that investigated the writing productions of native speakers of English.

Nevertheless, such a finding does not mean that the student-informants already arrived at a [nearly] advanced English proficiency. Moreover, the results of the structure and lexical bundle match-up procedure show that a majority of the lexical bundles used by the student-informants in developing their essays are completely obscure or cannot be clearly understood because a quarter of all the bundles were categorized as 'other expressions.' Consequently, most of them did not fit into any of the grammatical types forwarded by Biber et al. (1999). This would mean that the informants of the study do not have enough knowledge of English lexical bundles, and this might be attributed to their limited exposure to such frequently occurring word expressions. The students should, therefore, familiarize themselves with these English lexical bundles, and they should value the importance of these word combinations as "basic building blocks for constructing […] written discourse" (Biber & Conrad, 1999, p. 188).

While this research has contributed a lot to the existing knowledge of lexical bundles, further studies can be undertaken on the usage of recurrent word combinations, especially in international settings. Using a larger corpus, it would be interesting to investigate and compare the lexical bundles found in the academic corpus to the ones identified in previous studies conducted by foreigners so as to determine if students in Philippine colleges and universities employ the same word combinations in like manner. Moreover, a survey may be conducted among the student-informants to ascertain if they are cognizant of the use of English lexical bundles and the implications of such for academic writing.

References

Anthony, L. (2007). Antconc 3.2.1w: Freeware corpus analysis toolkit. Retrieved from http://www.antlab.sci.waseda.ac.jp/

Beng, C., & Keong, Y. (2014). A corpus study of structural types of lexical bundles in MUET reading texts. 3L: *The Southeast Asian Journal of English Language Studies, 20*(2), 127-140.

Biber, D. (2006). *University language: A corpus-based study of spoken and written registers.* Benjamin.

Biber, D., & Conrad, S. (1999). Lexical bundles in conversations and academic prose. In H. Hasselgard, & S. Oksefjell (Eds.), *Out of corpora: Studies in honor of Stig Johansson* (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*, 371-405.

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *The Longman grammar of spoken and written English*. Longman.

_____

Bowker, L., & Pearson, J. (2002). *Working with specialized language: A practical guide to using corpora.* Routledge.

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

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

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

Fuster-Marquez, M. (2014). Lexical bundles and phrase frames in the language of hotel websites. *English Text Construction, 7*(1), 84-121.

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

Kim, Y. (2009). Korean lexical bundles in conversation and academic texts. *Corpora, 4*, 135-165.

Nesi, H., & Basturkmen, H. (2006). Lexical bundles and discourse signaling in academic lectures. *International Journal of Corpus Linguistics, 11*, 283-304.

Pierini, P. (2009). Adjectives in tourism English on the web: A corpus-based study. *Círculo de Lingüística Aplicada a la Comunicación, 40*, 93-116.

Rafiee, M., & Keihaniyan, M. (2013). A comparative analysis of lexical bundles in journalistic writing in English and Persian: A contrastive linguistic perspective. *International Journal of Foreign Language Teaching and Research, 1*(2), 37-44.

Salazar, D.J.L. (2010). Lexical bundles in Philippine and British English. *Philippine Journal of Linguistics, 41*, 94-109.

Salazar, D.J.L. (2011). Lexical bundles in scientific English: A corpus-based study of native and non-native writing (Doctoral dissertation). Universitat de Barcelona. Retrieved from http://www.tdx.cat/bitstream/handle/10803/52083/DJLS_DISSERTATION.pdf?sequence=1