

## **An Analysis of Lexical Bundles in Research Article Abstracts by Iranian and Native English-speaking Authors of Applied Linguistics Articles**

Hesamoddin Shahriari Ahmadi, Behzad Ghonsooly and Azar Hosseini Fatemi

*Ferdowsi University of Mashhad, Iran*

### **Biodata**

Hesamoddin Shahriari Ahmadi is an assistant professor of applied linguistics at the department of English, Ferdowsi University of Mashhad. His chief research interests include second language writing, EAP and corpus linguistics.

Behzad Ghonsooly is a professor in applied linguistics. He teaches at the department of English, Ferdowsi University of Mashhad. His research interests are ESP and language testing.

Azar Hosseini Fatemi is an assistant professor in TEFL and the head of the department of English at Ferdowsi University of Mashhad. Her areas of interest include issues in second language teaching and learning.

### **Abstract**

Given the importance of effective communication among members of various academic disciplines, the research article has gained immense popularity among linguists, hoping to inform academic writing instructors and advanced EFL/ESL writers with the results of their

analyses. This study hopes to achieve a similar goal by analyzing research article abstracts in terms of the frequency and function of lexical bundles. A corpus of 200 research article abstracts by Iranian authors in the field of applied linguistics was compared to a similar corpus of abstracts by native English-speaking authors. The results reveal that Iranian authors use more 4-word lexical bundles in their writing compared to their native speaker counterparts and often use these bundles for achieving specific genre moves. A linguistic analysis of the bundles also shows that the lexical bundles found in the Iranian corpus included more clausal elements and subordination, while native speaker bundles were more phrasal in nature.

**Keywords:** lexical bundles, multiword expressions, corpus linguistics, research articles, abstracts

## 1. Introduction

University students, coming from both native- and foreign-language backgrounds, are expected to use the English language in ways differing from that which they had encountered during their high-school education or EFL courses. The process of adapting to a hitherto unfamiliar register can pose difficulties for students. For the EFL student, in particular, these problems are compounded by the additional complexities involved in mastering the language itself. In most cases, graduate students are required to write and publish academic research articles without having received the necessary training for the task. As a result of this shortcoming, it is important for researchers and practitioners in the field of EAP to investigate the features of academic writing and the variables contributing to the successful composition of research articles acceptable to members of their respective scientific communities.

Pawley and Syder (1983) maintain that the knowledge of a body of lexicalized sentence stems is what distinguishes highly proficient speakers of a language from less advanced learners. In addition to improving the quality of second language production, formulaic sequences have been shown to provide second language learners with a facilitated means of becoming communicative (Schmitt & Carter, 2004). That is to say, through the use of prefabricated lexical patterns, speakers and writers of a second language can perform their intended linguistic functions in a quick and easy way, and more easily integrate themselves into a peer group. In Iran, English is the language of communication for a number of university majors, especially the field of English language studies. This also holds true for students seeking to improve their academic writing proficiency; that is, graduate students for whom English is not a native language could potentially make use of prefabricated lexical sequences in order to attain higher levels of proficiency and to gain membership in the academic discourse community of their field by publishing their research findings. However, it is necessary for researchers to determine how advanced writers of a field are currently making use of formulaic strings in their writing, and also how these lexical sequences differ from those used by professional, published writers.

In Iranian universities, students of Applied Linguistics use English as the predominant language in textbooks, classroom lectures and discussions. Furthermore, university entrance exams include general language proficiency items, which guarantee the admission of students who are more advanced in terms of their general English language proficiency. However, despite their advanced level of communicative skills in English, students still struggle when it comes to assignments involving academic writing. Therefore, it should be determined whether using formulaic sequences could improve the academic writing quality of these students. One

possible way of answering this question is by comparing texts by advanced Iranian writers of Applied Linguistics with internationally-published writers of the same field, and considering how these two groups differ in their use of prefabricated lexical sequences. Assuming there is a continuum of proficiency in academic writing, it would be extremely useful for us to locate where advanced learners currently stand along the continuum and what the gaps which need to be bridged are so that they could achieve more efficient writing skills. This is by no means a simple task, because differences between the written output of advanced Iranian writers and the norms established within the register cannot be referred to as mistakes or errors, and should rather be viewed as deviations from the norm. These deviations are very difficult to single out by relying solely on one's intuition. Corpus-driven approaches to linguistic analysis are one way researchers can circumvent this problem. Instead of relying on intuitions regarding the nature of language, the corpus-driven approach to linguistic analysis relies on empirical data to answer language-related questions. The applications of corpus-driven research are numerous, and for these reasons, in the present study we have opted for such an approach in order to fulfill our primary objectives, which are described below.

The present study has two main objectives. First of all, using a corpus-driven methodology, it seeks to determine whether formulaic language, in the form of lexical bundles, does in fact play a role in the academic writing of a group of advanced Iranian L2 writers of English. This aim is achieved through the analysis of a corpus of Applied Linguistics research articles by Iranian writers. The second aim of the study is to determine how the identified lexical bundles differ from those employed by internationally published writers, most of whom use English as their native language. For this purpose, a parallel corpus consisting of research articles by native English speakers appearing in internationally-

accredited journals of Applied Linguistics was analyzed and the discovered bundles were compared with those found in the writing of Iranian writers in terms of frequency, structure and function. The findings of this study will inform us about how native and advanced non-native writers of a specific register differ in their use of pre-fabricated lexical strings. These results would be of use to academic writing instructors who wish to help their students write texts of higher quality by filling in the existing gaps between their current state of writing and that of professional, published writers. In the following section of this paper, the impetus underlying the use of corpus methodology, as well as the nature and definition of lexical bundles along with a summary of studies conducted on this unit, will be discussed in detail.

## **2. Review of related literature**

### **2.1 Lexical bundles**

In studying multiword expressions, some scholars have chosen a purely frequency-based approach (Altenberg, 1998; Butler, 1997; Biber et al., 1999). Biber and Conrad (1999) make use of a frequency-driven approach, analyzing the most frequently recurring sequences of words or extended collocations. They chose the term 'lexical bundles' to refer to combinations of words that occur repeatedly with a relatively high frequency within a given register. Unlike idioms that are non-compositional, lexical bundles are semantically transparent and, in most cases, their meaning can be understood from their components. Another difference between idioms and lexical bundles is that the latter are far more frequent in discourse. However, lexical bundles are not always complete in structure. Following this landmark study, the lexical bundle analysis framework was adopted by other researchers, such

as Partington and Morley (2004), Hyland (2008), Crossley et al. (2011), and Wei and Lei (2011).

Lexical bundles have also been used for comparing texts written by learners and authors coming from different language backgrounds. Cortes (2004) compared students' use of lexical bundles with patterns of use among published authors in the field of history and biology. The results of this study showed that students did not use lexical bundles as frequently as professional authors of the field, and in instances where the bundles were actually employed no correspondence was found between the patterns of use. Wei and Lei (2011) compared four-word lexical bundles across a corpus of doctoral dissertations by advanced Chinese EFL writers and published research articles by professional writers. They, however, found that advanced learners generally used bundles more frequently and with greater range. In another cross-linguistic comparative study, Chen and Baker (2010) used an automated frequency-driven method to compare lexical bundles in a corpus of published academic texts and a corpus of student academic writing. Learners were also observed to generally use a narrower range of lexical bundles, and at times overuse certain expressions hardly seen in the corpus of professional academic writers.

## **2.2 Classification of lexical bundles**

Biber (2006) identifies three types of lexical bundles. The first type includes verb fragments, beginning with a subject pronoun followed by a verb phrase (e.g. *I'm not going to*), a verb phrase, without a pronoun (e.g. *is going to be*), and a question fragment (e.g. *does that make sense*). Bundles beginning with discourse markers followed by a verb phrase (e.g. *you know it was*) and a question fragment (e.g. *I mean don't you*) are also categorized under this category.

The second type of bundle is characterized by verb phrase elements incorporating dependent clause fragments. These often include a complementizer following a main clause (e.g. *I don't know why*) or a WH-word presenting a dependent clause (*what I want to*). Dependent clause fragments starting with a complementizer or subordinator (e.g. *that it is a, as we will see*) also form Type 2 bundles. Finally, the third category involves phrasal components (noun phrase components often ending in a post-modifier, e.g. *the way in which*), prepositional phrase component with modifiers (*at the end of*) or incorporating comparative expressions (*as well as the*). Academic prose mostly incorporates phrasal lexical bundles (Biber, 2006).

### **2.3 Previous studies on lexical bundles**

Lexical bundle research has been carried out in both L1 and L2. In L1, studies have mostly identified bundles, describing their patterns of use across different registers (Biber et al., 1999; Biber & Conrad, 1999) and determining the discourse functions served by bundles in different texts (Cortes, 2004). Previous studies have also attempted to compare native and non-native English speakers and writers in terms of the characteristics of the lexical bundles they use. One group of studies has shown that there appears to be a difference between the two groups in terms of the overall number of bundles used (Erman, 2009; Howarth, 1998; Adel & Erman, 2012). A second group of studies shows that in addition to the difference in frequency, native and non-native users of English also differ with regard to the variety of bundles they use in their writing (Granger, 1998; Lewis, 2009). De Cock (2000) found that L2 users of English generally lacked awareness when it came to more common, yet less salient, L2 bundles, and often relied on L1 transfer to make up for their unawareness. L1 transfer occurred either through the modification or avoidance of forms which did not have an L1

equivalent. In the case of constructions where there was no match between L1 and L2, students commonly misused the L2 form. On the other hand, L2 users exhibited a tendency towards overusing those set of constructions with shared L1 equivalents.

Lexical bundles have been shown to frequently appear in academic registers. The frequency with which bundles are seen, however, depends on the size of the corpus and the number of words constituting the bundle. Biber et al. (1999) found that in academic texts, bundles consisting of three words occurred over 60,000 times and 4-word bundles appeared more than 5,000 times per million words. Although most words in a given text did not appear in recurrent combinations, almost 21% of the academic sub-corpus of the Longman Corpus of Spoken and Written English occurred in frequent bundles, some occurring at over 200 times per million words.

Considering the important role of lexical bundles in academic writing and the familiarity of competent writers of this discourse with a variety of these units (Hyland, 2008), attempts will now be made to compare and contrast instances of academic writing composed by Iranian authors writing in English with that of their native-English-speaking counterparts in terms of the lexical bundles employed and their respective frequency. This study will compare the two groups of writers within the register of applied linguistics research article abstracts. Through this attempt, we hope to reveal the frequency and category of bundles characterizing article abstracts by both groups of academics. Research carried out on lexical bundles in academic writing have commonly analyzed the research article in its entirety and have rarely zoomed in on a discipline in particular, let alone a single section of the article (i.e. the abstract). The findings, in addition to providing insights into the nature of how lexical bundles are used by authors from different linguistic backgrounds, could potentially guide



academic writing instructors on what forms of prefabricated patterns and phrases to highlight in their syllabi and in the course of their classes.

### 3. Method

#### 3.1 The corpus

The corpus used in this study consisted of 400 research article abstracts from the field of applied linguistics. The corpus was equally divided into two sub-corpora. The first sub-corpus consisted of 200 abstracts written by authors for whom English was a native language (NS), and the second sub-corpus was comprised of the same number of abstracts by Iranian authors, writing in English as their second language (INNS). The sample of texts in the NS corpus was randomly selected from four established journals, namely *Applied Linguistics*, *TESOL Quarterly*, *English for Specific Purposes*, and *Journal of Second Language Writing*. The second group of abstracts were taken from papers published in accredited journals published inside Iran; these journals were *Journal of Teaching Language Skills*, *Iranian Journal of Applied Linguistics*, *Iranian Journal of Applied Language Studies*, and *Research in Foreign Languages*. An equal proportion of article abstracts were taken from each of the journals specified above. Table 1 below shows the details of the corpus used in this paper.

Table 1: Details of sub-corpora used in the analysis

Journals	(NS)	Journals	(INNS)
<i>Applied Linguistics</i>	50	<i>Journal of Teaching Language Skills</i>	50
<i>TESOL Quarterly</i>	50	<i>Iranian Journal of Applied Linguistics</i>	50
<i>ESP Journal</i>	50	<i>Iranian Journal of Applied</i>	50

<i>Journal of Second Language Writing</i>	50	<i>Language Studies Research in Foreign Languages</i>	50
---	----	---	----

The NS sub-corpus consisted of 34,471 word tokens and 4,509 word types (the type-token ratio for this sub-corpus was 0.13), while the INNS corpus of abstracts consisted of 32,311 word tokens and 4,399 word types (the type-token ratio for this sub-corpus was also 0.13).

### 3.2 Procedure

The compiled corpus was analyzed using AntConc 3.3.0, a freeware concordance program developed by Laurence Anthony at the Center for English Language Education in Science and Engineering (CELESE), Waseda University (Japan). The N-gram function of this software was used to identify existing lexical bundles in each corpus through an empirical analysis. Our adopted definition of lexical bundles is borrowed from Biber et al. (1999), who define them as the most recurrent multi-word sequences in a given register. The study focused on units consisting of 4-word sequences. According to Hyland (2008), 4-word bundles are more common than 5-word bundles and usually represent more clear structures and functions than 3-word bundles, but he also admits that the cut-off points for bundles is quite arbitrary. The frequency cut-off used for the identification of bundles in this study was a minimum of 5 attestations. Since the two corpora in the analysis were of the same size (each including 200 texts), there was no need for norming. This criterion is believed to result in more bundles and to better suit the exploratory nature of this study. In order to avoid idiosyncratic expressions by individual authors, it was agreed that a sequence would have to appear in at least 5 different texts to be included into the analysis.

The lexical bundles found during the course of this investigation often do not represent complete structural units. Rather, they are mostly bridging elements, linking two structural units (i.e., phrases or clauses) to each other. The relatively high frequency with which these bundles recur in the corpus reveals that they are most likely stored and used as pre-fabricated linguistic patterns. As a result, even though the identified bundles did not always constitute a whole structural unit, they could be used as an index for determining the degree to which authors use the formulaic principle to form their utterances.

#### **4. Results and Discussion**

##### **4.1 Extracted lexical bundles**

The analysis of N-grams of 4-word clusters generated a list of lexical bundles. The NS corpus contained 18 bundles, while the INNS corpus consisted of 41 bundles. The most frequently recurring bundle in the NS corpus was *this paper reports on*, occurring 12 times. Following this, the most frequent bundles in this sub-corpus were *paper reports on a*, *in the use of*, and *the extent to which*, with 9, 8, and 8 occurrences, respectively. On the other hand, in the INNS corpus, the most frequently repeated lexical bundle was *the results of the*, occurring 17 times, followed by *the purpose of this*, *as a foreign language*, and *English as a foreign*, each with 11 instances. None of the top-ten most frequent lexical bundles in either one of the two corpora was shared by the other. In fact, the only lexical bundles common between the two lists were *of English as a*, *as a foreign language*, and *the results of the*. The first two of the common bundles constitute various four-word combinations in the phrase *English as a foreign language*, which is an expectedly common phrase in the field of applied linguistics. The bundle *the results of the* shows that the authors in both groups used this bundle to discuss the

findings of their study in brief in their abstracts. Some of the bundles in the INNS sub-corpus (e.g. of Iranian EFL learners) were obviously repeated as a result of the common subject areas investigated by Iranian authors. Table 2 below provides a list of the 4-word lexical bundles along with their frequency and the number of texts in which they appeared. The lexical bundles common between the two corpora have been presented in bold.

Table 2: Most frequent 3- and 4-word lexical bundles found in the two sub-corpora

Corpus of Abstracts by Iranian Authors (INNS)		Corpus of Abstracts by Native Speakers (NS)			
Bundles	Frequency	No. of Texts	Bundles	Frequency	No. of Texts
<b>The results of the</b>	17	14	This paper reports on	12	12
The purpose of this	11	9	paper reports on a	9	9
<b>as a foreign language</b>	11	8	in the use of	8	7
English as a foreign	11	8	the extent to which	8	8
findings of the study	10	6	the use of English	8	6
significant difference between the	9	9	the ways in which	7	5
The results indicated that	9	8	a wide range of	6	5
of the present study	8	8	in terms of the	6	5
of this study was	8	8	in the field of	6	6
The findings of the	8	6	the results of a	6	6
of Iranian EFL learners	7	6	<b>The results of the</b>	6	6
of the study revealed	7	7	<b>as a foreign language</b>	5	7
purpose of this study	7	6	discussed in terms of	5	5
results of the study	7	7	It was found that	5	5
this study was to	7	7	<b>of English as a</b>	5	5
The analysis of the	7	5	reports on a study	5	5
aim of this study	6	6	the role of the	5	5
revealed that there was	6	6	This study investigated the	5	5
study was to investigate	6	6			
that there was a	6	6			
the Iranian EFL learners	6	5			
The purpose of the	6	6			
results of this study	6	5			
The results showed that	6	5			
as a result of	5	5			
divided into two groups	5	5			

investigate the effect of	5	5		
is an attempt to	5	5		
no significant difference between	5	5		
<b>of English as a</b>	5	5		
of the study are	5	5		
of this study is	5	5		
that there is a	5	5		
The results revealed that	5	5		
the study are discussed	5	5		
the study revealed that	5	5		
This paper reports the	5	5		
this study is to	5	5		
to investigate the effect	5	5		
was to investigate the	5	5		
was used to analyze	5	5		

#### 4.2 Structural classification of lexical bundles

The identified lexical bundles were consequently classified into one of three main categories introduced by Biber et al. (1999). As is characteristic of academic prose and the research article register, in both sub-corpora, the majority of bundles were phrasal. From the list of bundles found in the INNS corpus, 19 were phrasal; on the other hand, 12 phrasal bundles were observed in the NS corpus of abstracts. The second most common type of lexical bundle in the INNS corpus was the verb phrase element followed by a dependent clause (12 instances); this category of lexical bundles was only found once in the NS sub-corpus. Finally, 5 lexical bundles in the NS corpus and 7 in the INNS corpus were verb phrase fragments, either with or without a preceding subject. Table 3 below shows the number and percentage of the different categories of bundles in the analyzed corpora.

Table 3. Types and categories of lexical bundles in the two corpora

	Verb Phrase Fragments		Verb Phrase Elements + Dependent Clauses		Phrasal Components	
	N	Percentage	N	Percentage	N	Percentage
NS Corpus	5	27.7%	1	5.5%	12	66.6%
INNS Corpus	7	46.3%	12	29.2%	19	17%

### 4.3 Functional classification of lexical bundles

Bhatia (1994) suggests that research article abstracts provide readers with information about various aspects of the research article, including what the author did during the study; how it was done; what the author discovered; and what conclusions were drawn from the findings. Just like other sections of the research article, the abstract consists of a number of steps and moves. Bhatia (1994) found that the genre of research article was comprised of four moves: *Introducing the purpose*, *Describing the methodology*, *Summarizing the results*, and *Presenting the conclusions*. A functional analysis of the lexical bundles shows that in both Iranian (16) and native-speaker (4) abstracts, the largest number of bundles was used to introduce the purpose of the study. A number of examples of such bundles in context are provided below:

**The purpose of this** study was to explore the factors that influenced... (NS Corpus)

**The purpose of this** paper is to show that research orientation... (NNS Corpus)

The second most common function for the lexical bundles found was summarizing the results and presenting the conclusions. Sixteen bundles in the INNS corpus and four bundles in the

NS corpus were used for these functions. These two functions are very similar in nature and tend to overlap in terms of the linguistic features and lexical bundles they possess. For instance, the exact same lexical bundle could be found serving both functions in two different abstract corpora. In the examples below, the lexical bundle *findings of the study* is used to summarize the results in the first sentence and to present the conclusions of the study in the second.

The **findings of the study** revealed no significant relationship between... (NNS Corpus)

The **findings of the study** prove that language planning is ideologized... (NNS Corpus)

Finally, the least number of lexical bundles in either corpus was devoted to describing the methodology of the study. The NS and INNS sub-corpora had 2 and 1 bundles devoted to this function, respectively. The bundle *in terms of the* is commonly used for presenting the conclusions of the study, as follows:

...instances of interactional trouble are identified and discussed **in terms of the** teachers' elaboration of some routine features... (NS Corpus)

...These findings are discussed **in terms of the** participants' local concerns... (NS Corpus)

The greater number of lexical bundles found in the INNS corpus confirms the findings of other studies which have also discovered that advanced writers in English as a second/foreign language tend to overuse lexical bundles in their academic writing compared to professional authors for whom English is a native language (Cortes, 2004; Hyland, 2008; Wei, 2007; Pang, 2009; and Wei & Lei, 2011). Various studies have shown that some non-native speakers tend to use certain formulaic sequences frequently and repetitively, because they view them as

reliable 'safety nets' which can be confidently used, especially at times of uncertainty (De Cock, 2000; Granger, 1998). In other words, lexical bundles could be viewed as a form of compensation strategy for non-native writers of English.

Another interesting finding of this study was that most of the lexical bundles found in the INNS corpus were clearly used for achieving one of the abstract moves, while a considerable portion of bundles in the NS corpus (e.g. *in the use of*, *the extent to which*, *the ways in which*, etc.) did not belong to any particular move and could generally be used for various purposes. This gives further plausibility to the previous claim that the use of lexical bundles in the writing of advanced learners may be a learning strategy, by means of which learners can achieve the move structure of any given genre more easily and with greater clarity. These lexical bundles function as 'genre frames', serving a dual purpose: First, they assist writers in more easily fulfilling genre expectations (e.g. introducing the purpose of the study, summarizing the results) and adding to the clarity of the passage, just as discourse markers enhance a reader's comprehension of a text. Second, they reduce the cognitive load imposed on writers by providing ready-made sequences stored as prefabricated units of language.

The classification of lexical bundles based on Biber et al.'s (1999) taxonomy revealed that Iranian writers used considerably more verb phrase fragments and verb phrase elements followed by dependent clauses. Biber et al. (2011) claim that clausal subordination is more commonly observed in academic conversation compared to academic writing, which tends to include more noun phrase constituents and complex phrasal structures. According to these findings, Iranian EFL writers, despite their high level of proficiency and expertise as specialists in the field of applied linguistics, tend to write in an academically conversational



style, depending heavily on clausal elements and subordination. This finding is quite natural, since academic writing, with its reliance on extended noun phrases and limited use of verbs, is quite counterintuitive and can only be seen in the academic writing register, which strives for greater brevity and density, conveying more information in fewer words.

## **5. Conclusion**

This study investigated the use of 4-word lexical bundles in research article abstracts written by native English speakers and Iranian EFL writers in the field of applied linguistics. Through a corpus-driven approach to linguistic analysis, the two groups of abstracts were compared in terms of the number, type, and communicative function of lexical bundles with reference to the move-structure of research article abstracts. Results showed that advanced writers tend to use more multi-word expressions in their writing, and also more commonly employ those expressions for serving specific genre expectations. Writing instructors can use the findings of this study to further emphasize the role of lexical bundles for their students, explaining to them the various ways these prefabricated language patterns can be used to achieve moves and steps, consequently aiding target readers in understanding their passage and heightening their chances of publishing their work in accredited journals. Future studies can identify more lexical bundles used in different academic disciplines and construct an inventory of lexical bundles, used as genre frames. Such an inventory would most certainly be of great interest to academic writing instructors and advanced EFL writers alike.

The more prominent use of clausal fragments and subordination in abstracts by Iranian authors is interesting, since it is different from native speaker norms in academic writing and more closely resembles academic speech. Other written registers of English, such as fiction

and journalistic texts, are different in that they use more extensive subordinate clauses and rely less on noun phrases and nominalizations. The fact that abstracts by Iranian authors in this study did not highlight noun phrase structures as in the case of native speaker abstracts could possibly mean that Iranian writers, despite their relatively high level of proficiency as published authors in the field of applied linguistics, were still on a developmental path, moving from a general formal writing style towards the academic writing register.

The findings also have implications for writing instructors active in the field of English for Academic Purposes (EAP). According to frequency models of language learning, phraseological units are learned probabilistically and as a result of strengthening associations among co-occurring words (Ellis, 2002). The more frequently a learner is exposed to particular bundles, the higher the chances are for that learner to produce the bundles in speech or writing. This linear relationship between exposure and output is often times affected by L1 background. If a bundle shares an equivalent in the L1, the probability of its use will increase. As previously discussed, the proclivity towards extended noun phrase elements in the academic writing register is unnatural in that it is rarely seen in other registers. It is also very uncommon for other languages (in this case, Farsi) to follow such a trend; hence, learners, even those at advanced proficiency levels, find it counterintuitive and avoid using bundles representing noun phrase elements. As a result, it is recommended that EAP teachers and academic writing instructors create necessary awareness and improve the quality of academic writing by exposing learners to high-frequency phrasal bundles used by native speakers.

The relationship between the lexical bundles found in this study and their functioning as genre-frames could also be of great instructional value. Since bundles such as *the purpose of this* are immediately associated with a particular move in abstracts, we could also

consequently claim that readers encountering this bundle in an abstract would also more readily identify its aim. Therefore, it is clearly an advantage for any writer to make use of high-frequency bundles belonging to a specific genre move. Lexical bundles, as other multiword expressions, are known to be stored holistically. This results in quicker accessibility and easier processing (Vogel Sosa & MacFarlane, 2002). Not only will this allow the writer to communicate the intended message more efficiently and with greater ease, but it would also make it easier for the reader to grasp the writer's intention. Future studies could possibly look into the frequency and function of lexical bundles used by learners/writers from other L1 backgrounds or other parts of the research article. The techniques through which lexical bundles can be taught or brought to learners' attention would also be an interesting topic for further investigation.

## References

- Adel, A., & Erman, B. (2012). Recurrent word combinations in academic writing by native and non-native speakers of English: A lexical bundle approach. *English for Specific Purposes* 31, 81-92.
- Altenberg, B. (1998). On the phraseology of spoken English: The evidence of recurrent word combinations. In A. Cowie (Ed.), *Phraseology: Theory, analysis and applications* (pp. 101-122). Oxford: Oxford University Press.
- Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.

- Baker, P., & Chen, Y. (2010). Lexical bundles in L1 and L2 academic writing. *Language, Learning and Technology, 14*(2), 30-49.
- Biber, D. (2004). Lexical bundles in academic speech and writing. In B. Lewandowska-Tomaszyk (Ed.), *Practical applications in language and computers—PALC 2003* (pp. 165-178). Frankfurt am Main: Peter Lang.
- Biber, D. (2006). *University language*. Amsterdam: John Benjamins.
- Biber, D., & Conrad, S. (1999). Lexical bundles in conversation and academic prose. In H. Hasselgard & S. Oksefjell (Eds.), *Out of corpora: Studies in honor of Stig Johansson* (pp. 181-189). Amsterdam: Rodopi.
- Biber, D., Gray, B., & Ponpoon, K. (2011). Should we use characteristics of conversation to measure grammatical complexity in L2 writing development? *TESOL Quarterly, 45*(1), 5-35.
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. London: Longman.
- Butler, C. S. (1997). Repeated word combinations in spoken and written text: Some implications for Functional Grammar. In C. S. Butler, J. H. Connolly, R. A. Gatward, & R. M. Vismans (Eds.), *A fund of ideas: Recent developments in functional grammar* (pp. 60-77). Amsterdam: IFOTT, University of Amsterdam.
- Cortes, V. (2004). Lexical bundles in published and student disciplinary writing: Examples from history and biology. *English for Specific Purposes, 23*, 397-423.
- Crossley, S., & Salsbury, T. L. (2011). The development of lexical bundle accuracy and production in English second language speakers. *International Review of Applied Linguistics in Language Teaching, 49*(1), 1-26.

- Deane, P., & Quinlan, T. (2010). What automated analyses of corpora can tell us about students' writing skills. *Journal of Writing Research*, 2(2), 151-177.
- De Cock, S. (2000). Repetitive phrasal chunkiness and advanced EFL speech and writing. In C. Mair & M. Hundt (Eds.), *Corpus linguistics and linguistic theory* (pp. 51-68). Amsterdam: Rodopi.
- Diniz, L., & Moran, K. (2005). Corpus-based tools for efficient writing instruction. *Essential Teacher*, 2(3), 36-39.
- Ellis, N. C. (2002). Frequency effects in language processing. *Studies in Language Acquisition*, 24(2), 143-188.
- Ellis, N. C. (2008). Phraseology: The periphery and the heart of language. In F. Meunier & S. Granger (Eds.), *Phraseology in foreign language learning and teaching* (pp. 1-13). Amsterdam: John Benjamins.
- Erman, B. (2009). Formulaic language from a learner perspective: What the learner needs to know. In B. Corrigan, H. Quali, E. Moravcsik, & K. Wheatley (Eds.), *Formulaic language* (pp. 27-50). Amsterdam: John Benjamins.
- Flowerdew, L. (1996). Concordancing in language learning. In M. C. Pennington (Ed.), *The power of CALL* (pp. 97-114). Houston, TX: Athelstan.
- Granger, S. (1998). Prefabricated patterns in advanced EFL writing: Collocations and formulae. In A. P. Cowie (Ed.), *Phraseology: An interdisciplinary perspective* (pp. 145-160). Amsterdam: John Benjamins.
- Gries, S. T. (2008). Dispersions and adjusted frequencies in corpora. *International Journal of Corpus Linguistics*, 13(4), 403-437.