

A Comparative Study of the Use of Lexical Bundles in Academic Writing by Iraqi and Expert Authors

Assistant Lecturer

Hayder S. Alkhafaji

hayder_alkhafaji@iunajaf.edu.iq

Assistant Lecturer

Ahmed Thaer Abdulzahra

ahmed_thaer94@iunajaf.edu.iq

Assistant Lecturer

Aqeel Mery Alabedi

aqeelmiri@alkafeel.edu.iq

Islamic University of Najaf

دراسة مقارنة لاستخدام الحزم المعجمية في الكتابة الأكاديمية من قبل

المؤلفين العراقيين والخبراء

المدرس المساعد

حيدر صباح الخفاجي

المدرس المساعد

أحمد ثائر عبد الزهرة

المدرس المساعد

عقيل ميري محمد

الجامعة الإسلامية - النجف الأشرف

Abstract:-

The research piece has gained a lot of attention from linguists who want to share their findings with academic writing instructors and advanced EFL/ESL writers, given the necessity of good communication among members of diverse academic fields. Research article abstracts can be studied in terms of the frequency and use of lexical bundles, this study seeks to achieve a similar goal. Linguists have been fascinated by common word combinations and multiword formulations over the past two decades. The current study focuses on a specific sort of such word pairings known as lexical bundles, which are a set of three or more words that frequently occur in the same register (Biber, 2004). There is a correlation between the frequency of lexical bundles in published research paper abstracts written by Iraqi and Expert writers in two different fields of study. This study's corpus includes over 2000 research papers authored by two sets of writers from the disciplines of engineering and humanities. In the 300,000-word corpus, the analysis showed significant a list of 4-word lexical bundles that appeared at least 30 times per million words (150,000 words for each of the two main corpora by Iraqi and Expert Writers). Furthermore, for each of the bundles in the list, each of the two corpora was searched separately. Iraqi writers use fewer lexical bundles than Expert writers, according to the data. To look at it another way, Iraqi authors used a smaller variety of lexical bundles than Expert authors. However, in the Iraqi corpus of abstracts, lexical bundles had a higher frequency of use. The study found that Iraqi academic writers frequently used a small number of lexical bundles, while experts used a wider range of bundles with a lower token frequency. It was found that the structural and functional use patterns of Iraqi and Expert writers' lexical bundle categories were similar despite the differences between the two main corpora.

Keywords: lexical bundles, IR corpus, EX corpus, engineering, humanities discipline, frequency, structural and functional categories

الملخص:

حظيت هذه الورقة البحثية باهتمام كبير من اللغويين الذين يرغبون في مشاركة نتائجهم مع مدربي الكتابة الأكاديمية وكتاب اللغة الإنجليزية كلغة أجنبية / اللغة الإنجليزية كلغة ثانية المتقدمين، نظراً لضرورة التواصل الجيد بين أعضاء المجالات الأكاديمية المتنوعة. يمكن دراسة ملخصات المقالات البحثية من حيث تكرار واستخدام الحزم المعجمية، وتسعى هذه الدراسة إلى تحقيق هدف مماثل. كان اللغويون مفتونين بتركيبات الكلمات الشائعة والصيغ المتعددة الكلمات على مدى العقدين الماضيين. تركز الدراسة الحالية على نوع معين من أزواج الكلمات المعروفة باسم الحزم المعجمية، وهي مجموعة من ثلاث كلمات أو أكثر تحدث بشكل متكرر في نفس السجل (Biber, 2004). هناك علاقة ارتباط بين تكرار الحزم المعجمية في ملخصات الأوراق البحثية المنشورة التي كتبها كتاب عراقيون وخبراء في مجالين دراسيين مختلفين. تشمل مجموعة هذه الدراسة على أكثر من ١٠٠٠ ورقة بحثية كتبها مجموعتان من الكتاب من تخصصات الهندسة والعلوم الإنسانية. في مجموعة الكلمات المكونة من ٣٠٠٠٠٠ كلمة، أظهر التحليل قائمة مهمة من الحزم المعجمية المكونة من ٤ كلمات والتي ظهرت ٣٠ مرة على الأقل لكل مليون كلمة (١٥٠.٠٠٠) كلمة لكل من المجموعتين الرئيسيتين لكتاب عراقيين وخبراء. علاوة على ذلك، تم البحث في كل مجموعة من المجموعتين على حدة. وبحسب البيانات، يستخدم الكتاب العراقيون حزم معجمية أقل من الكتاب الخبراء. للنظر إلى الأمر بطريقة أخرى، استخدم المؤلفون العراقيون مجموعة أصغر من الحزم المعجمية من المؤلفين الخبراء. ومع ذلك، في مجموعة الملخصات العراقية، كانت الحزم المعجمية ذات تواتر أعلى في الاستخدام. وجدت الدراسة أن الكتاب الأكاديميين العراقيين استخدموا بشكل متكرر عدداً صغيراً من الحزم المعجمية، بينما استخدم الخبراء نطاقاً أوسع من الحزم بتردد رمزي أقل. وقد وجد أن أنماط الاستخدام الهيكلية والوظيفية لفئات الحزمة المعجمية للكتاب العراقيين والخبراء كانت متشابهة على الرغم من الاختلافات بين المجموعتين الرئيسيتين.

الكلمات المفتاحية: الحزم المعجمية، الهندسة، العلوم الإنسانية، الانضباط، التكرار، الفئات الهيكلية والوظيفية

1. Introduction

There has been a change in the study of language corpora and formulaic sequences as a result of technological and computer breakthroughs during the previous two decades. Researchers used to analyze formulaic language only on the basis of intuition, grouping a number of phrases together based on theoretical rationales. However, advances in computer technology over the last two decades have enabled academics to seek for these recurring units of language in a more systematic manner. The development of a new form of these units, referred to as lexical bundles, is one of the consequences of this new systematic method to finding formulaic phrases in the language. Lexical bundles, according to Biber, Johansson, Leech, Conrad, and Finegan, are "the most frequently encountered sequences of words" in a given register (1999). (Biber, 2006, p. 134).

Furthermore, lexical bundles differ from other forms of formulaic language in a few respects. Bundles, for example, do not frequently transmit idiomatic meanings in their role in the language (Biber, Conrad, & Cortes, 2004). On the other hand, in expressions like on the other hand, the extent to which, and in the area of the words that make up these bundles retain their own meaning after the phrase is formed. Furthermore, while there are some notable exceptions, the majority of lexical bundles do not constitute structurally complete units. Cortes (2013) identified a wide variety of bundles with six to nine words in her corpus study of research paper beginnings. Some of these bundles were structurally complete, contradicting the widely held belief that lexical bundles are incomplete units.

Biber and colleagues (2004) developed three primary categories to help better understand lexical bundles: stance expressions, discourse organizing devices, and referential phrases. They discovered a distinction between the ways in which various functional categories of bundles were used in written and spoken modes. Academic writing is more likely to employ referential bundles than attitude bundles, which are more prevalent in casual conversation. By contrast, bundles of discourse structuring are less prevalent.

The primary goal of this study is to determine the frequency with which lexical bundles are used in academic writing by Iraqi researchers who have published abstracts in national/Iraqi scientific

journals. Additionally, the structure and function of these bundles will be examined and compared to a parallel corpus of research article abstracts written by expert authors who have published their work in internationally recognized journals in similar fields. This comparison will contribute to discovering lexical bundles that are overused, underutilized, or misused by Iraqi researchers of scientific research publications written in English. A similar comparison may be conducted between academic disciplines to examine which academic discipline's usage of lexical bundles more closely resembles patterns of use by professional authors in the same subjects. The research will also look at lexical bundles that are exclusive to a group of writers. For instance, lexical bundles occurring exclusively in engineering abstracts produced by Expert Authors and not in engineering abstracts written by Iraqi writers are detected. A list like this might help us better understand how Iraqi academic researchers employ lexical bundles, in addition to those that are used by both groups.

Consequently, researchers and scholars from all countries want to increase their scientific knowledge and improve the quality of their higher education, and one method to do so is to publish research articles in international publications. The fact that English is widely regarded as a lingua franca encourages academics and researchers to publish their findings in this language. Because of this, publishing a research paper is one of the criteria for completing a doctorate or master's degree.

Professional Iraqi Researchers, academic writing tutors, and syllabus/material creators will benefit from the outcomes of this study. Previous researches into the use of lexical bundles by scientists from various L1 backgrounds has been useful in not only improving the process of academic writing instruction for that specific group of scientists and researchers, but also in gaining a better understanding of the process of second/foreign language academic writing and its development. Materials that address the specific challenges and difficulties experienced by authors from a particular L1 background might be created based on the findings of such studies. Curricula may be customized to their specific requirements and fields of study in this way. As a result, it will assist instructors and researchers in implementing the most effective instructional approaches and tactics for pupils in order to reach and surpass their learning objectives.

1.2. Statement of the Problem

It's crucial to understand how professional writers from various L1 backgrounds employ multiword phrases and formulaic sequences in their academic work. Few, if any, studies have looked at how professional Iraqi writers employ lexical bundles and how these bundles affect their writing output across several fields, particularly in engineering, and the humanities.

For a variety of academic disciplines in Iraq, English is the predominant language of publishing. To enhance the quality of writing in these journals (i.e., International Academic Iraqi journals) and to make it possible for Iraqi academics to publish their work in reputable, international, peer-reviewed journals, we must first understand the language used by professional Iraqi writers who currently write and publish their work in Iraqi national journals. As researchers also need to learn how this writing differs from academic English used in international publications in the same discipline. The latter group of articles exemplifies effective academic writing and may be used as a standard for others aspiring to the same level of achievement (i.e., gaining wider coverage of their work through publication in international, academic journals). Due to the lack of Arabic-language publications, the study will concentrate on some disciplines where most professional Iraqi authors are expected to publish their research in English. The use of formulaic sequences and multiword phrases in academic writing by Iraqi professional writers can be better understood by observing how lexical bundles are used in academic writing by Iraqi writers.

1.3. Research Questions

The purpose of this study is to address the following research questions:

1. What are the most frequently occurring lexical bundles in Iraqi research paper abstracts in comparison to experts in Engineering and Humanities?
2. When comparing Iraqi researchers to experts in the fields of Engineering and Humanities, What are the main prominent structural and functional classifications of lexical bundles?

2. Theoretical Background

2.1 Lexical Bundles

Biber et al. (1999) coined the term "lexical bundle" in "Longman Grammar of Spoken and Written English." According to him, lexical bundles are frequently occurring groups of three or more words that occur in both the written and spoken word (Biber et al., 1999). In other words, lexical clusters of three to four words that occur frequently in spoken or written language are neither complete grammatical units nor idiomatic in nature (Biber, 2006). Biber (2007, p. 264) defined lexical bundles as "frequently occurring sequences of words [...] lexical bundles are frequently structurally incomplete and lack idiomatic meaning, but they serve critical discourse functions in both spoken and written texts." In other words, lexical bundles are simply groups of words that are frequently used together in both written and spoken language.

Hyland (2008, p. 41) describes lexical bundles as "expanding collocations which seem more frequently than anticipated, reshaping meanings and going to contribute to our sense of coherence in a text." However, learners tend to use these bundles in their written or spoken register, despite the occasional overuse and misuse of lexical bundles, because it helps them to produce meaningful text or speech. Using lexical bundles, according to Haswel (1991), can lead to the development of a more effective writer or speaker.

Authors and language learners from around the world have used lexical bundles to compare their writing styles. Students' use of lexical bundles was compared to the use of history and biology terms by published authors in Cortes (2004). Lexical bundles were less frequently used by students than by professionals in this field, and there was no correlation between the two groups' usage. They looked at the lexical bundles in doctoral dissertations by advanced Chinese EFL writers and professional research articles (Wei & Lei, 2011). Using bundles was more common and varied among more advanced students, according to the researchers. The lexical bundles found in published academic texts and student essays were compared by Chen and Baker (2010) in another cross-linguistic study. In addition, students tended to use fewer lexical bundles and overuse expressions that are rarely used in academic writing.

2.2 Functional Taxonomy of Lexical Bundles

According to Biber et al. (2006), the functions of lexical bundles can be divided into three broad groups. In the next section, we'll look at how to use stance expressions, discourse organizers, and referential expressions.

2.2.1 Stance Expressions

Bundle stances explain and clarify attitudes or certainty assessments that include some intents in the context of receiving information. Both epistemic and attitude/modality can be categorized (Biber, 2006). First-person pronouns (I) and impersonality (It's obvious that) can be used to identify stance bundles, which can also be used to convey meaning that isn't directly tied to the authors or speakers.

Epistemic stance bundles convey the degrees of certainty or probability. In other words, they express the meaning of certainty, uncertainty, and probability/possibility. These bundles are primarily personal, especially in the spoken registers, which commonly demonstrate ambiguity, as in the example that's difficult to tell, but again. Irreverent irony is conveyed by the fact that the Russian Revolution declared itself to be Marxist in aims and character, but it happened contrary to Marxist historical logic. Similarly, impersonal stance bundles frequently convey a message of certainty. In (Biber's 2006 book, p. 140), you can find this information.

Bundles of attitude/modality stances convey the various ways in which speakers can convey their position on an action. These bundles are categorized as follows: obligation/directive, intention/prediction, desire, and capability (Biber, 2006). The first of these categories, obligation/directive, expresses a personal stance differently than the personal bundles discussed previously, in that it uses a second-person pronoun (you) rather than a first-person pronoun (I). By utilizing these bundles (i.e., obligation/directive bundles), the speakers encourage the addressee to take action (Biber, 2006). This is demonstrated clearly in the following example; all that remains is for you to work on it (Biber, 2006, p. 140).

The second category is intention/prediction bundles. Bundles like these convey the speaker's intentions or predictions for the future,

and thus, it is concluded that they are primarily a statement of personal preference as in the example, however, for the time being, we'll focus on those that are generated and found to be beneficial (Biber, 2006, p. 141). And sometimes, these express impersonal stance by conveying predictions of the speaker for the coming events or actions which do not involve the speakers' desire. Imagine the following example: if you involve a 20% return on investment, the investment's net present value will be zero (Biber, 2006, p. 141).

The other two categories of attitude/modality stance bundles are desire and ability. Desire bundles always convey a speaker's personal stance by expressing the speaker's wishes and desires, as in this example, so I may avoid seeing her face to face in order to avoid delivering bad news to her (Biber, 2006, p. 141). Finally, there are capability bundles. These bundles express the speaker's capability and define the skills and tasks that the speaker should perform.

2.2.2 Discourse organizing bundles

There are two primary functions of discourse organizing bundles: introducing and focusing on a topic, and elaborating on this topic (Biber, 2006). The purpose of the introduction/focus is to inform the student that a new topic will be discussed in this example. Please take a look at the exercise we'll be doing. Here's an example of what I'm talking about: That's (Biber 2006, p. 142)

Discourse organizing bundles also play a major role in clarifying or elaborating on a topic. This example (Biber, 2006, p. 144) shows how these bundles can be used to provide additional explanation and clarification:

We know that if the project is in the same line business as the firm's other projects [...] then high standalone risk translates into high corporate risk [...]. On the other hand, if the project is not in the same line business, then it is possible that the correlation may below....

2.2.3 Referential bundles

The final type of lexical bundles are those that refer to other words. An important part of a unique feature, the items included in these bundles can exist. Reference bundles are classified into four major subcategories: identification, imprecision indicators, attribute specification, and time/place/text reference.

References in the imprecision subcategory serve as a placeholder for more precise references, or they can be used to point to additional sources of similar information (Biber, 2006). In the following examples, I believe we have somewhere around six weeks left in class., there are firms that conduct evaluations and similar services, and it's self-evident (Biber, 2006, p. 145)

The second subcategory is a specification of attributes. These bundles identify or specify the following nouns that come after them, giving the nouns specific attributes. Some of these bundles identify amounts or quantities. Other bundles identify the size and form of the nouns that come after them (Biber, 2006). As in these examples, does it result in significant wealth creation? No. It generates a small amount of wealth. These figures illustrate the size of Russia's ethnological community. (2006) (Biber, p. 145)

The third subcategory encompasses time, place, and textual reference. These bundles make reference to a specific time, place, or location within the text in order to emphasize a particular quality. For instance, the more proactive real estate agents were in the following examples, as illustrated in Figure 4.4, the more houses they sold... As a result, you must record the sale of the asset at the end of the year. In the United States, children are not formally employed in farm work (Biber, 2006, p.146)

2.3 Structural Taxonomy of Lexical Bundles

Additionally, lexical bundles can be categorized according to the structure of their structural correlates. Biber et al. (2004) develop a structural typology from their examination of lexical bundles in university registers. In summary, despite the fact that they are not complete structural units, lexical bundles have strong grammatical correlations (Biber et al, 2004). The following tables summarize the structural characteristics of lexical bundles.

Table 2.1. Structural Taxonomy of Lexical Bundles (Biber et al, 2004)

Lexical bundles that incorporate verb phrase fragments	
Structure type	Example
1st/2nd person pronoun + VP fragment 3rd person pronoun + VP fragment Discourse marker + VP fragment	I'm going to, you don't have to It's going to be, that was one of the I mean you know, you know it was

Verb phrase (with a non-passive verb)	Is going to be, is one of the
Verb phrase with a non-passive verb	Is based on them, can be used to
Yes/no question fragments	Are you going to, do you want to
WH question fragments	What do you think, how many of you

Table 2.3. Structural Taxonomy of Lexical Bundles (Biber et al, 2004)

Lexical bundles that incorporate noun phrases and prepositional phrases fragments	
Structure type	Example
Noun phrase with of-phrase fragments	One of the things, the end of the
Noun phrase with other post-modifier fragments	A little bit about, those of you who
Other noun phrase expressions	And stuff like that, a little bit more
Prepositional phrase expressions	Of the things that, at the end of
Comparative expressions	As far as the, greater than or equal

3. Method

3.1 The corpus

The corpus for this study contained research paper abstracts from the disciplines of Human sciences and Engineering, authored by professional published researchers and L2-Professional Iraqi researchers, and published between 2010 and 2020. For the overall quantity of words, the two corpora were nearly identical (300,000 words in total and around 150,000 words in each corpus). In fact, the abstracts in the IR (Iraqi) corpus are often shorter than those in the EX (Expert) corpus articles, furthermore, the IR corpus has more articles overall. The IR corpus included a random sample of texts from Iraqi Academic Scientific Journals, whereas the EX corpus of research article abstracts came from two well-known journals, Oxford Journals, and Cambridge Journals.

Based on the following considerations, the journals chosen for the IR corpus are thought to be substantially equivalent to the EX corpus. High-ranking universities in Iraq produce the journals in the IR corpus. Furthermore, we are more likely to represent the language created by professional Iraqi authors in an Iraqi academic environment by picking L2 articles from Iraqi Academic Scientific Journals, which are published in Iraq. As a result, a comparison of these two corpora reveals the parallels and contrasts in the written discourse of L1-English academic professional writers against L2-Iraqi academic professional authors. The two corpora are also large

enough to analyze patterns of lexical bundles or recurring word combinations.

3.2 Procedure

AntConc version 3.4.4 was used in this investigative process to extract four-word lexical bundles. Numerous studies have examined lexical bundles of varying sizes, with sometimes random selection criteria; however, Hyland (2008) asserted that 4-word lexical bundles are more prevalent than 5-word bundles and generally indicate more distinct structure and function than 3-word bundles. Because they are so common and often indicate recurrent collocations, lexical bundles of only two words were also omitted from the analysis. As a guideline, 30 occurrences per million words was used to identify each and every bundle. Additionally, a bundle must appear in at least five texts in order to be considered for the final list. As a safeguard against the unusual examples of usage provided by some writers, this precaution was taken.

To identify bundles, Biber et al (2004) emphasize that frequency is critical. A pattern in frequency data "... must be explained," he said. (See page 376 for more information). Additionally, frequency "reflects the extent to which a sequence of words is stored and used as a prefabricated chunk, with higher frequency sequences being more likely to be stored as unanalyzed chunks than lower frequency sequences" in the study of word sequences (Biber et al, p.376). The frequency with which these bundles appear in a million words is arbitrary, but this study used the 30 occurrences per million rule. The study of lexical bundles and their frequency was conducted using AntConc version 3.4.4.

Only four-word multiword phrases or bundles were examined in this study, as in earlier studies on lexical bundles (e.g., Cortes, 2002). When all of the tests were evaluated, the algorithm found all of the bundles that appeared 30 times in over 600,000 words. Biber et al. stipulate that MWEs must occur at least ten times per million words in a register and be recognized in at least five texts (1999). Only four-word sequences were analyzed because five- and six-word utterances are uncommon, and three-word bundles "may be viewed as an extended collocational linkage" (Biber et al, 1999).

Immediately following the creation of the bundles, the structural and functional relationships between them were investigated in great depth. Both the structural and functional taxonomies were created by Biber et al. (1999), with the functional taxonomy created by Biber, Conrad, and Cortes (2003). (2003; 2004). This is illustrated in Table 3.1, which shows the rubric that was used for the structural and functional analysis of bundles.

Table 3.1. Functional taxonomy of the lexical bundles (Biber et al. 2004, pp. 384-388)

Category	Sub-categories	Examples
I. Stance Expressions		
	A. Epistemic Stance	
	Personal	I thought it was
	Impersonal	Are more likely to
	B. Attitudinal/Modality Stance	
	b.1. Desire	
	Personal	I don't want to
	b.2. Obligation/directive	
	Personal	I want you to
	Impersonal	It is necessary to
	b.3. Intention/prediction	
	Personal	I'm not going to
	Impersonal	Are going to be
	b.4. Ability	
	Personal	To be able to
	Impersonal	It is possible to
II. Discourse Organizers		
	A. Topic introduction/focus	If we look at
	B. Topic elaboration/clarification	On the other hand
III. Referential Expressions		
	A. Identification/focus	One of the most
	B. Imprecision	And things like that
	C. Specification of attributes	
	c.1. Quantity specification	Greater than or equal
	c.2. Tangible framing attributes	In the form of
	c.3. Intangible framing attributes	The nature of the
	D. Time/place/text reference	
	d.1. Place reference	In the United States
	d.2. Time reference	At the same time
	d.3. Text deixis	As shown in figure
	d.4. Multi-functional reference	At the end of
IV. Special conversational functions	A. Politeness	Thank you very much
	B. Simple inquiry	What are you doing
	C. Reporting	I said to him/her

4. Results and Discussion

The study's primary aim was to ascertain the most frequently used lexical bundles in academic research paper abstracts. The combined corpus of research paper abstracts, which contains over 300,000 words, was selected for this purpose because frequency is the major distinguishing attribute of lexical bundles. Both corpora, the EX corpus, and the IR corpus have about 150.000 words in each. The researchers will compare lexical bundles between Iraqi and Expert writers in this section of the study, in addition to determining the most common lexical bundles. To begin, the researchers will provide a list of bundles from the corpus, as well as their frequencies. Then we compare Iraqi lexical bundles to vocabulary bundles used by Expert authors in their abstracts to see how they are distinct and/or similar in the three subjects.

4.1. Lexical bundles in the IR and EX corpus from the discipline of Engineering

To a degree, the engineering discipline's most prevalent lexical bundles differed from those discovered in the preceding discipline. In other words, engineering writers frequently work with a variety of bundles. Table 4.7 summarizes the lexical bundles identified in the Iraqi and Expert Engineering disciplines.

Table 4.7. List of lexical bundles of IR and EX Engineering discipline

No.	Bundle of Iraqi Engineering discipline	Freq.	No.	Bundle of Expert Engineering discipline	Freq.
1	results showed that the	38	1	can be used to	22
2	was found that the	26	2	in this study we	15
3	results demonstrate that the	23	3	a wide range of	13
4	The findings revealed that	25	4	in the presence of	13
5	the results show that	24	5	in the absence of	13
6	the results of the	22	6	an important role in	13
7	it is found that	22	7	in this paper we	12
8	in the present work	21	8	at the same time	10

9	the performance of the	20	9	in an attempt to	10
10	has been used to	20	10	the development of a	10
11	in this paper the	20	11	as a result of	10
12	is one of the	20	12	of this study was	9
13	study the effect of	20	13	on the basis of	9
14	the effect of the	20	14	in the case of	9
15	in the range of	18	15	the use of the	9
16	to study the effect	17	16	for the development of	8
17	one of the most	16	17	it is possible to	7
18	of the most important	16	18	as well as the	7
19	as well as the	14	19	the aim of this	7
20	as a result of	13	20	this study was to	6
21	the aim of this	13	21	in the field of	6
22	on the other hand	12	22	in the context of	6
23	an increase in the	11	23	for the treatment of	6
24	in addition to the	10	24	in the present study	6
25	in the field of	9	25	is one of the	6
26	can be used to	9	26	in addition to the	5
27	a wide range of	9	27	the analysis of the	5
28	in this study the	9	28	one of the most	5
29	was found to be	9	29	on the other hand	5
30	one of the important	9	30	the role of the	5
31	the importance of the	8	31	was found to be	5
32	the aim of the	8	32	is based on the	4
33	in terms of the	8	33	the development of the	4
34	for the purpose of	8	34	the performance of the	4
35	is based on the	8	35	a large number of	4

36	of this paper is	8	36	this article	4
37	a result of the	7	37	examines the	4
38	this paper is to	7	38	this study's	4
39	of this study is	7	39	objective	3
40	in the form of	7	40	of this study is	3
41	it is possible to	7	41	the aim of the	3
42	the objective of this	6	42	results show that	3
43	throughout this study	6	43	the	3
44	the results of this	6	44	in the	3
45	while at the same time	6	45	development of	3
46	the role of the	6	46	the results show	3
47	this study is to	6	47	that	3
48	were found to be	5	48	is the most	3
49	objective of this study	5	49	common	3
50	in the presence of	4	50	has been used to	3
51	the relationship between the	4	51	in the form of	3
52	the use of the	4	52	in terms of the	3
53	on the basis of	3	53	it is argued that	3
54	the purpose of this	3	54	the results of this	3
55	in the development of	3	55	on the use of	3
56	the study of the	3	56	this study is to	2
57	in the case of	3	57	the importance of	2
58	the end of the	3	58	the	2
59	as one of the	3	59	the results of the	2
60	in the absence of	3	60	the purpose of	2
61	it was found that	2	61	this	2
				at the university	2
				of	2
				the effect of the	2
				in the range of	2
				methods a total of	2
				a result of the	2
				an increase in the	2
				in relation to the	2
				of the most	2
				important	2

62	the nature of the	2	62	results showed	2
63	in an attempt to	2	63	that the	2
64	the development of the	2	64	the nature of the	2
65	for the treatment of	2	64	the results	2
66	in this paper we	2	65	showed that	2
67	were included in the	2	65	study was carried	2
68	was carried out in	2	66	out	2
69	the analysis of the	2	66	was found that	2
70	the development of a	2	67	the	2
71	aim of this study	2	67	in this study the	2
72	this study was to	2	68	study the effect of	2
73	of this study was	2	69	to study the effect	2
74	aim of the study	2	70	in this paper the	1
75	in the context of	2	71	the end of the	1
76	study was carried out	2	72	the objective of	1
77	at the university of	2	73	this	1
78	an important role in	2	73	objective of this	1
79	in the light of	1	74	study	1
80	on the one hand	1	74	of this paper is	1
81	a large number of	1	75	in this article we	1
82	in relation to the	1	76	on the one hand	1
			77	this paper is to	1
			78	was carried out in	1

Iraqi writers were more likely to use the bundles than Expert writers. For instance, the bundle appears 38 times in the IR corpus, but only once in the EX corpus. Likewise, was discovered that occurs 26 times in the IR corpus but only once in the EX corpus. This demonstrates that Iraqi writing, in comparison to expert writing in the field of engineering, makes excessive use of these bundles. The most prevalent is the engineering discipline's unique lexical bundle (i.e., absent in the other disciplines).

4.2. Lexical bundles in the IR and EX corpus from the Humanities discipline

The two lists of IWs and EWs bundles found in the humanities discipline are shown in table 4.8.

Table 4.8. List of lexical bundles of IR and EX Humanities discipline

No.	Bundle of IR Humanities discipline	Freq.	No	Bundle of EX Humanities discipline	Freq.
1	as well as the	42	1	the international criminal court	39
2	one of the most	40	2	and crimes against humanity	30
3	of the most important	46	3	as well as the	34
4	is one of the	42	4	of crimes against humanity	28
5	the nature of the	38	5	at the same time	24
6	the city of Najaf	36	6	in the context of	24
7	the importance of the	31	7	on the basis of	24
8	as one of the	26	8	at the university of	24
9	as a result of	24	9	the role of the	24
10	in addition to the	22	10	one of the most	22
11	in the field of	20	11	it is argued that	18
12	the study of the	19	12	on the use of	18
13	for the purpose of	19	13	this article examines the	18
14	the importance of this	18	14	in the field of	12
15	the role of the	17	15	in the development of	11
16	at the same time	16	16	in the case of	10
17	in the light of	15	17	as one of the	9
18	on the other hand	12	18	the purpose of this	9
19	the aim of the	12	19	in the light of	9
20	in an attempt to	12	20	the relationship between the	8
21	the development of the	11	21	the development of	8

22	one of the important	10	22	is one of the	8
23	on the basis of	10	23	the nature of the	7
24	at the university of	10	24	in the form of	7
25	the end of the	9	25	the development of the	7
26	the use of the	8	26	the study of the	7
27	to find out the	8	27	the end of the	6
28	of the twentieth century	7	28	in this article we	6
29	in terms of the	7	29	in relation to the	5
30	on the one hand	6	30	the importance of the	5
31	a large number of	6	31	on the one hand	5
32	in relation to the	6	32	of the most important	5
33	the analysis of the	6	33	on the other hand	5
34	of this study is	6	34	a wide range of	4
35	the results of the	6	35	for the purpose of	4
36	in the development of	5	36	in an attempt to	4
37	is based on the	5	37	in the presence of	4
38	in the form of	5	38	is based on the	4
39	a result of the	5	39	the use of the	4
40	this study was to	4	40	for the development of	3
41	in the context of	4	41	in the absence of	3
42	an important role in	4	42	of this paper is	3
43	the relationship between the	4	43	of the twentieth century	3
44	for the development of	4	44	the aim of this	3
45	aim of the study	4	45	in addition to the	2
46	in this study the	3	46	in this paper we	2
47	the effect of the	3	47	this paper is to	2
48	the aim of this	3	48	as a result of	2

49	aim of this study	3	49	of this study is	2
50	results showed that the	3	50	this study is to	2
51	in the presence of	3	51	can be used to	1
52	the performance of the	3	52	an important role in	1
53	in the case of	2	53	a large number of	1
54	in this paper the	2	54	the analysis of the	1
55	in the absence of	2			
56	in this paper we	1			
57	the results of this	1			
58	this paper is to	1			
59	it was found that	1			
60	of this study was	1			
61	this study is to	1			
62	the purpose of this	1			
63	study the effect of	1			
64	to study the effect	1			
65	of this paper is	1			
66	an increase in the	1			
67	on the use of	1			
68	the development of a	1			

Comparing the lexical bundles found in the humanities with those found in the other discipline, it is possible to see the difference in how Iraqi authors utilize lexical bundles, Iraqi writers, on the other hand, use more bundles and more frequently than Expert writers. When examining the list of bundles discovered in the IR corpus, the first, as well as the most common, occurs 42 times, followed by one of the most common, which occurs 40 times, and one of the most important, which occurs 46 times. While both occur 34 times in the EX corpus and one of the most frequently occurs 22 times.

The first two lexical bundles in the EX corpus, the international criminal court occurs 39 times, and crimes against humanity occur 30

times, are unique lexical bundles found in Expert writing, in addition to their absences in the IR corpus or humanities discipline. Table 4.9. provide lists the distinct bundles encountered in the engineering discipline.

Table 4.9.
List of unique lexical bundles in the humanities discipline

No.	Unique bundle
1	and crimes against humanity
2	of crimes against humanity
3	the international criminal court

4.3. Structural Classifications of Lexical Bundles

The Biber et al. taxonomy was used to classify the lexical bundles in this study (2004). This taxonomy recognized each of the three major structural categories of lexical bundles: noun phrases, verb phrases, and dependent sentence fragments (Shahriari, 2017). The Structure of a lexical bundle refers to the grammatical pattern that it adopts or is embedded in; for instance, the end of is understood to be a bundle that incorporates a noun phrase fragment.

The noun phrase is the most prevalent structural category in the corpus, according to the total list of lexical bundles (61bundles, equal to 63.5 % of the overall bundles). The dependent clause (22 bundles, accounting for 22.9 % of all bundles) is the second most prevalent structural category, followed by the verb phrase category (13 bundles, equal to 13.5 % of the overall bundles). Figure 4 depicts how lexical bundles are distributed throughout the three major structural groups.

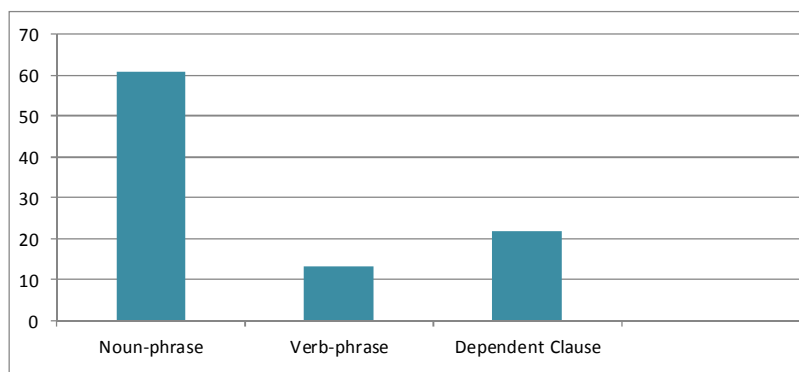


Figure 4: Lexical Bundles' Structural Classifications in the Entire Corpus

Noun-phrase bundles are the most prevalent structural type in the IR and EX corpora, with the IR corpus containing 55 bundles (equivalent to 66.2 % bundles in the IR corpus) and the EX corpus containing 57 noun phrase bundles (equal to 58.7 % of the total bundles in the EX corpus). The dependent sentence is the second most frequent structural type, with 14 bundles in the IR corpus (equivalent to 16.8 % of all bundles in the IR corpus); by contrast, the EX corpus contains 18 bundles (equal to 18.5 % of the total bundles in the EX corpus).

Finally, the verb-paraphrase is the least prevalent structural type. The IR corpus has ten bundles of this sort (equivalent to 12% of the total bundles in the IR corpus), whereas the EX corpus contains thirteen bundles (equal to 13.4 % of the total bundles of the EX corpus). Figure 5 depicts how lexical bundles are distributed throughout the three major structural groups.

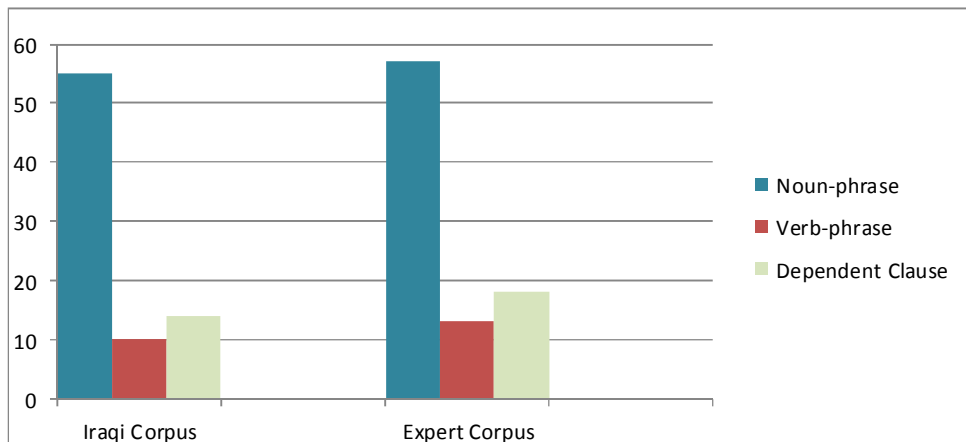


Figure 5: The number of lexical bundles found in the IR and EX corpora for each structural category.

There are 51 noun phrase bundles in the IR engineering dataset, while there are 50 in the EX engineering corpus (equal to 63.2 % of the total bundles in the EX engineering corpus). The next structural category is the dependent clause, which is the second most common. 21 bundles were found in the IR engineering corpus and 18 bundles were found in Ex engineering corpus (representing 25.3 % bundles found in the IR engineering corpus) (equal to 22.7 % of the

total bundles in the EX engineering corpus). There were only ten verb phrase bundles found in both the IR and EX engineering corpora, making it the least common structural type (equal to 12 % of the total bundles in the IR engineering corpus and equal to 12.6 % of the total bundles in the EX engineering corpus). FIGURE 7 depicts the distribution of word groups across the three structural categories.

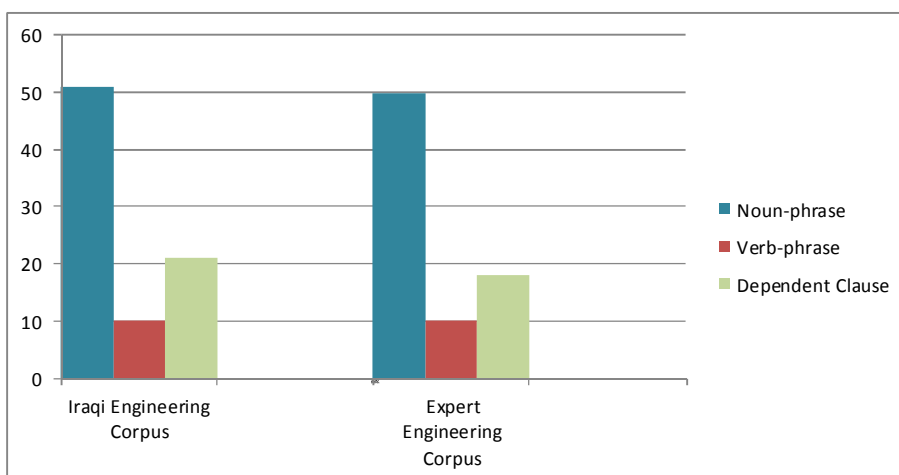


Figure 7. The IR and EX engineering corpus lexical bundles for each structural category

To begin, 50 noun phrase bundles were identified in the IR humanities corpus (equivalent to 72.4 % of all bundles in the corpus), whereas 41 bundles were identified in the EX humanities corpus (equal to 73.2% of the total bundles in the EX humanities corpus). The dependent sentence is the second most prevalent structural type, with 14 bundles discovered in the IR humanities dataset (20.2 % of all bundles in the IR humanities corpus) and 10 in the EX humanities corpus (equal to 17.8% of the total bundles in the EX humanities corpus). The verb phrase is the least frequent structural category, with only three detected in the IR and EX humanities corpora (equal to 4.3 % of the total bundles in the IR humanities corpus, 5.3% of the total bundles in the EX humanities corpus). The distribution of lexical bundles within each of the three major structural groups is depicted in Figure 8.

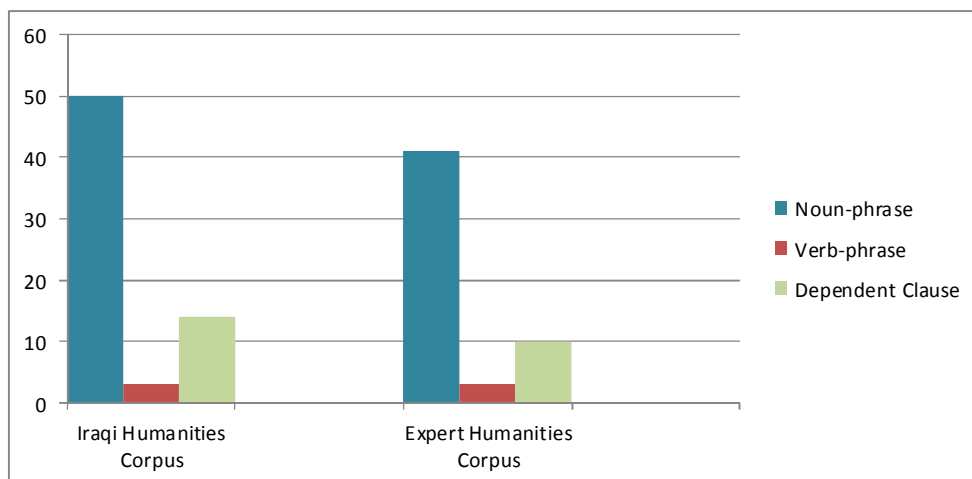


Figure 8: The IR and EX Humanities corpus lexical bundles for each structural category.

4.4. Functional Classifications of Lexical Bundles

According to the functional taxonomy of lexical bundles, the three most important functions of lexical bundles are stance, discourse organization, and reference. The writer's stance reveals his or her attitude, judgment, point of view, proposition or ability in terms of certainty or ambiguity. Bundles of geographical or chronological references are referred to as "reference bundles," while bundles of discourse organization are referred to as "discourse organizing bundles" (Biber, et al, 2004). The subclasses of lexical bundles are summarized in Table 3.1.

The referential bundle (which accounted for 55.3 % of all bundles in the combined IR and EX corpora) was the most prevalent functional category in the overall list of lexical bundles (i.e., in the combined IR and EX corpora) (equal to 35.3 % of the total bundles in the IR and EX corpus). The IR corpus contained 50 referential bundles (equivalent to 60.2 % bundles in the IR corpus), 23 discourse structuring bundles (equivalent to 27.7 % bundles in the IR corpus), and only six stance bundles (equivalent to 7.2 % bundles in the IR corpus). 56 referential bundles were detected in the EX corpus (equivalent to 68.3 % bundles in the I EX corpus), 28 discourse

structuring bundles were detected in the EX corpus (equivalent to 29 % bundles in the EX corpus), and only three attitude bundles were detected (equal to 5.2 % of the total bundles in the EX corpus). The distribution of lexical bundles across the three major structural groups is shown in Figure 9.

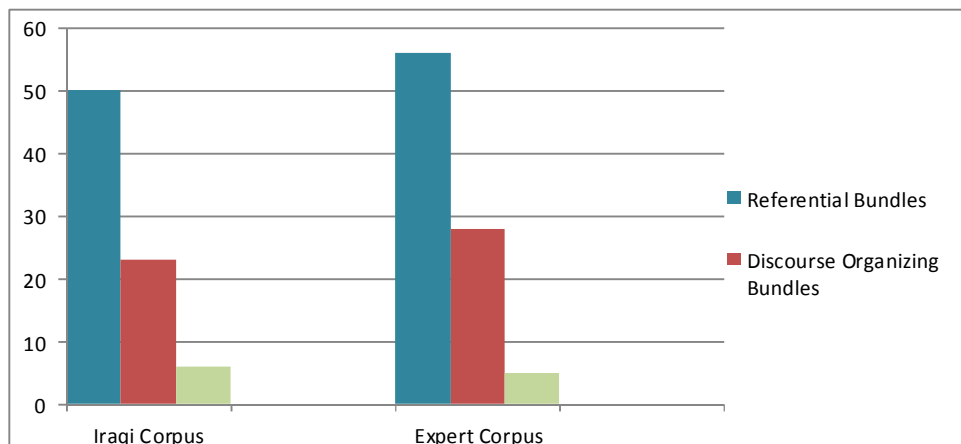


Figure 9: IR and EX lexical bundles for each functional category

Referential bundles (52 bundles, or 63.4 % of the bundles in the IR engineering corpus) are the most frequently occurring structural category in the IR engineering corpus, whereas the EX engineering corpus contained 49 bundles (equal to 62.8 % of the total bundles in the EX engineering corpus). The IR engineering corpus contained 24 discourse organizing bundles (equivalent to 29.2 % of all bundles in the IR engineering corpus), whereas the EX engineering corpus contained 23 such bundles (equal to 29.4 % of the total bundles in the EX engineering corpus). Finally, the IR identified only five bundles of stances, whereas the EX engineering corpus identified four bundles (equal to 6 % of the total bundles in the IR engineering corpus and equal to 5 % of the total bundles in the EX engineering corpus). As illustrated in Figure 11, there are three major structural groups into which lexical bundles are classified.

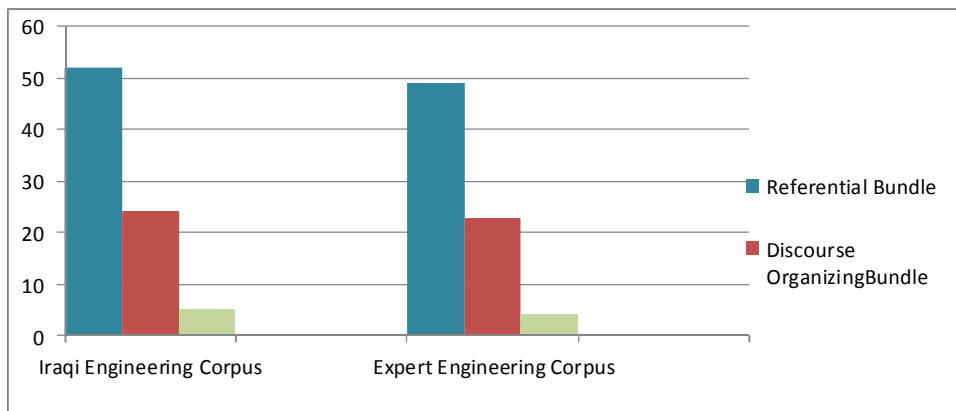


Figure 11: The IR and EX engineering corpus lexical bundles for each functional category

The IR humanities corpus contains 70.5 % noun phrase bundles, whereas the EX humanities corpus contains only 36 bundles (equal to 66.6% of the total bundles in the EX humanities corpus). Discourse organizers were the second most prevalent functional category, accounting for 19 bundles (or 27.9 % of all bundles in the IR humanities corpus) and 14 bundles in the EX humanities corpus (equal to 25.9 % of the total bundles of the EX humanities corpus). Only one bundle is devoted to the functional category of attitude in the EX humanities corpus (equal to 1.4% of the total bundles in the EX humanities corpus). The distribution of lexical bundles across the three major structural groups is depicted in Figure 12.

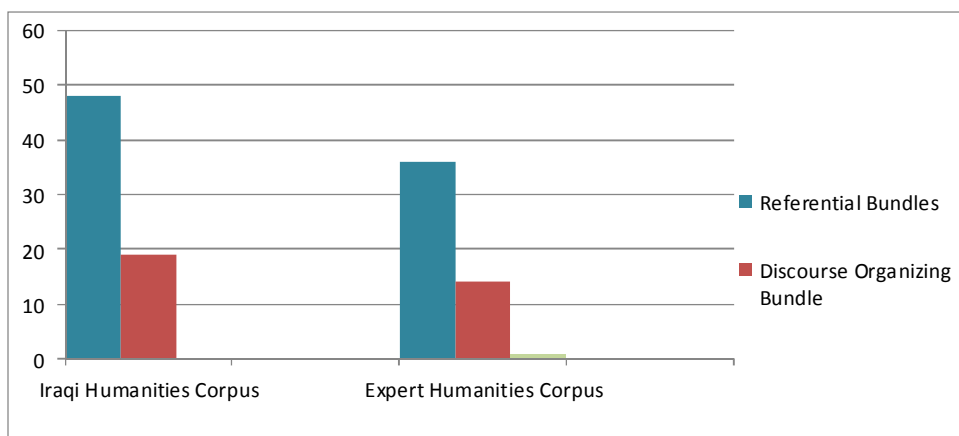


Figure 12: The IR and EX humanities corpus lexical bundles for each functional category

5. Conclusion

The primary goal of this study was to determine the number of lexical bundles used by Iraqi and Expert writers in research abstracts and to compare them across engineering and humanities disciplines. A computer program found four-word bundles in the Iraqi and Expert-writer corpora. The lexical bundles were reorganized structurally and functionally. We discovered some structural similarities between the writing styles of Iraqi and Expert authors in the IR and EX corpora. The IR corpus contained 50 referential bundles and 23 discourse organizers; the EX dataset contained 56 referential bundles, 28 discourse organizers, and five stance bundles.

The current study's findings can be understood in terms of their instructional implications. To begin, teaching vocabulary has been found to be useful in boosting learners' writing skills, and "Vocabulary is not simply memorizing individual words, but phraseology" (Ghani, 2016, p. 64). While lexical bundles are not always idiomatic or structurally complete, they are critical components of discourse (Biber & Barbieri, 2007) and should be emphasized during the academic writing training process. Second, the use of lexical bundles can improve writing for publication purposes, as well as assist and assist learners in effectively communicating ideas and using the language (Ranjbar, Pazhakh & Gorjian, 2012).

From a pedagogical standpoint, the current study's findings may be useful for textbook and syllabus designers in providing learners with more effective teaching materials. The authors of teaching materials, particularly those focused on improving English for Publication Purposes, will be able to utilize the findings of this study to choose relevant four-word phrases (i.e., lexical bundles) for inclusion in the syllabus.

The current investigation gives preliminary data on the quantity, structural pattern, and functional categories of lexical bundles in two fields (i.e., engineering, and humanities). Future research might benefit from examining patterns of lexical bundle use, including both structural and functional patterns, across other fields to find disciplinary variance in lexical bundle use. Instead of relying just on the abstract, future researchers should analyze the complete research piece, including the introduction, methodology, and

conclusion. Although this study has made an essential initial step in researching lexical bundles, especially in the two fields of the current study and adding to the existing understanding of lexical bundles, more research is needed to clarify a variety of lexical bundle forms and discourse roles.

Finally, no statistical significance tests were used to examine patterns of use between the two groups of writers in the current study. This was because a descriptive overview was attempted as a first investigation into the usage of lexical bundles. Future studies can look into this work and do statistical tests to see whether the changes found are significant.

References:

- Ädel, A. (2016). Salazar, Danica. Lexical Bundles in Native and Non-native Scientific Writing: Applying a Corpus-Based Study to Language Teaching. International Journal of Learner Corpus Research, 2(1), 125-129. <http://dx.doi.org/10.1075/ijlcr.2.1.06ade>
- Biber, D. (2006). University language (1st Ed.). Amsterdam: J. Benjamins..
- Biber, D. (2004). If you look at....: Lexical Bundles in University Teaching and Textbooks. Applied Linguistics, 25(3), 371-405. <http://dx.doi.org/10.1093/applin/25.3.371>
- Biber, D. & Gray, B. (2010). Challenging stereotypes about academic writing: Complexity, elaboration, explicitness. Journal Of English For Academic Purposes, 9(1), 2-20. <http://dx.doi.org/10.1016/j.jeap.2010.01.001>
- Biber, D. & Barbieri, F. (2007). Lexical bundles in university spoken and written registers. English for Specific Purposes, 26(3), 263-286. <http://dx.doi.org/10.1016/j.esp.2006.08.003>
- Biber, D. & Reppen, R. The Cambridge handbook of English corpus linguistics (1st Ed.)
- Byrd, P., & Coxhead, A. (2010). On the other hand: Lexical bundles in academic writing and in the teaching of EAP. University of Sydney Papers in TESOL.
- Biber, D. (2009). A corpus-driven approach to formulaic language in English: Multi-word patterns in speech and writing. International Journal of Corpus Linguistics, 14(3), 275-311. <http://dx.doi.org/10.1075/ijcl.14.3.08bib>
- Chen, Y. H., and Baker, P. (2010). Lexical bundles in L1 and L2 academic writing. Language Learning and Technology, 14: 30-49.

- Cortes, V. (2004). Lexical bundles in published and student disciplinary writing: Examples from history and biology. *English For Specific Purposes*, 23(4), 397-423. <http://dx.doi.org/10.1016/j.esp.2003.12.001>
- Cortes, V. (2008). A comparative analysis of lexical bundles in academic history writing in English and Spanish. *Corpora*, 3(1), 43-57. <http://dx.doi.org/10.3366/e1749503208000063>
- De Cock, S. (1998). A Recurrent Word Combination Approach to the Study of Formulae in the Speech of Native and Non-Native Speakers of English. *International Journal Of Corpus Linguistics*, 3(1), 59-80. <http://dx.doi.org/10.1075/ijcl.3.1.04dec>
- Dontcheva-Navratilova, O. (2012). Lexical Bundles in Academic Texts by Non-native Speakers. *Brno Studies In English*, 38(2), 37-58. <http://dx.doi.org/10.5817/bse2012-2-3>
- Granger, S. & Meunier, F. (2008). *Phraseology* (1st Ed.). Amsterdam: John Benjamins Pub
- Hyland, K. (2008). As can be seen: Lexical bundles and disciplinary variation. *English for Specific Purposes*, 27(1), 4-21. <http://dx.doi.org/10.1016/j.esp.2007.06.001>
- Hyland, K. (2008). Academic clusters: text patterning in published and postgraduate writing. *International Journal of Applied Linguistics*, 18(1), 41-62. doi:10.1111/j.1473-4192.2008.00178.x
- Jalali, Z. & Moini, M. (2014). Structure of Lexical Bundles in Introduction Section of Medical Research Articles. *Procedia - Social and Behavioral Sciences*, 98, 719-726. <http://dx.doi.org/10.1016/j.sbspro.2014.03.473>
- Jafarova, L. (2016). Challenges in Translating Phraseological Units. *Theory and Practice in Language Studies*, 6(7), 1341. <http://dx.doi.org/10.17507/tpls.0607.02>
- Jalali, Z. & Moini, M. (2014). Structure of Lexical Bundles in Introduction Section of Medical Research Articles. *Procedia - Social and Behavioral Sciences*, 98, 719-726. <http://dx.doi.org/10.1016/j.sbspro.2014.03.473>
- Meunier, F. & Granger, S. (2008). *Phraseology in foreign language learning and teaching* (1st Ed.). Amsterdam: John Benjamins Publishing.
- Hewings, M. & Hewings, A. (2002). "It is interesting to note that..." a comparative study of anticipatory 'it' in student and published writing. *English for Specific Purposes*, 21(4), 367-383. [http://dx.doi.org/10.1016/s0889-4906\(01\)00016-3](http://dx.doi.org/10.1016/s0889-4906(01)00016-3)
- Nesi, Hilary & H. Basturkmen. (2006). Lexical bundles and discourse signaling in academic lectures. (2006). *Lexical Cohesion and Corpus Linguistics*, 11(3), 283-304. <http://dx.doi.org/10.1075/ijcl.11.3.04nes>

- Ranjbar, N., Pazhakh, A., & Gorjian, B. (2012). The Effect of Lexical Bundles on Iranian EFL Learners Linguistic Production Fluency. *International Education Studies*, 5(4). <http://dx.doi.org/10.5539/ies.v5n4p243>
- Omidian, T., Shahriari, H., & Ghonsooly, B. (2016). Evaluating the Pedagogic Value of Multi-Word Expressions Based on EFL Teachers' and Advanced Learners' Value Judgments. *TESOL Journal*. <http://dx.doi.org/10.1002/tesj.284>
- Staples, S., Egbert, J., Biber, D., & McClair, A. (2013). Formulaic sequences and EAP writing development: Lexical bundles in the TOEFL iBT writing section. *Journal of English for Academic Purposes*, 12(3), 214-225. <http://dx.doi.org/10.1016/j.jeap.2013.05.002>
- Shin, Y. (2016). Lexical Bundles in Native and Non-Native Scientific Writing: Applying a Corpus-based Study to Language Teaching. *English for Specific Purposes*, 41, 82-83. <http://dx.doi.org/10.1016/j.esp.2015.07.005>
- Pan, F., Reppen, R., & Biber, D. (2016). Comparing patterns of L1 versus L2 English academic professionals: Lexical bundles in Telecommunications research journals. *Journal of English for Academic Purposes*, 21, 60-71. <http://dx.doi.org/10.1016/j.jeap.2015.11.003>
- Yaoyu Wei, & Lei Lei, (2011). Lexical Bundles in the Academic Writing of Advanced Chinese EFL Learners. *RELJ Journal*, 42(2), 155-166. <http://dx.doi.org/10.1177/0033688211407295>

