Article Information

🖊**Article Type:** _Research Article_

🖊**This article has been checked for similarity.**

INVESTIGATING NATIVE AND NON-NATIVE AUTHORS' USE OF LEXICAL BUNDLES IN THE LITERATURE OF ELT ARTICLES DISCUSSION SECTION

Maryam Shahrokhi Shahraki[1]

Mahdi Astaraki[2]

Abstract

Lexical bundles in ELT articles are considered as the bases of coherent contexts. In line with some previous studies, the present article aimed to compare native and nonnative ELT articles in terms of four-word lexical bundles. In so doing, a corpus including 200 ELT articles were chosen from 12 academic journals. First, all lexical bundles were identified and classified structurally and functionally based on Biber et al. (1999) taxonomy and Hyland (2008a, 2008b) functional category of lexical bundles in academic text respectively. To analyze the corpora, Antconc software (version 3.3.2), proposed by Anthony (2012), was used. The results revealed that authors of both corpora made use of various types of four-word lexical bundles; however, there were significant differences between native and nonnative articles in terms of their use of four-word lexical bundles; Iranian authors used theses lexical bundles almost twice more than native authors. In terms of functions, it was found that text-oriented bundles were the most frequent lexical combinations used by native and nonnative authors. Considering the results and previous studies, it is inferred that employing different lexical bundles to show the significant parts of academic research articles can help to effective information delivery in academic writing.

Keywords: Corpus linguistics, ELT Research articles, Genre analysis, Lexical bundles

[1]Assistant Professor of linguistics, Department of Linguistics and Foreign Languages, Payame Noor University, Tehran, Iran. E-mail: shahrokhymaryam@pnu.ac.ir ORCID: 0000-0002-5063-8622

[2]MA in Teaching English as a Foreign Language, Farhangian University, Tehran, Iran. E-mail: mahdii.astarakii1372@gmail.com ORCID: 0000-0003-0386-9257

Introduction

The discussion sections of the articles can be used as the mirror for the ELT writers` special discourse, style, and the genre and this MA thesis tries to show the differences (if there are any) between the discourse, style and genre of native and nonnative writers according to the lexical bundles as a kind of mark for the ELT writes discourse which they have used in their discussion sections. Analyzing these sections and the ELT writers' discourse in these parts of the ELT articles can be kind of criteria for checking the differences that can be among different ELT writers and this may have a kind of impact on conveying the message and information of the written articles on their readers due to different (easy, difficult, familiar and unfamiliar ones) lexical bundles that have been used in those sections.

Lexical bundles in the ELT articles have concerned the investigators not only due to their dominance in language but also for the reason they been discovered to be the bases of coherent context, and close discipline, genre, and register opposes (Biber et al. 2004) and hence, they are significant for the creation and comprehension of writing in the college (Biber 2006).

Moreover, specialists such as language teachers, English professionals and scientists would like to have adequate reading skills in order to get enough information through journal articles. In all, it would be beneficial for learners and teachers to be familiar with academic papers so that they would have a less difficult time in reading academic papers in the English language or even in writing them. In ELT courses, one area that learners have to be familiar with is words that appear in journal articles. In addition, Ward and Litman (2007) believed that by teaching students to process words in groups (i.e. noun groups of two, three or even four words); their reading can be improved to a higher level. So, it can be said that multi-word expressions are important components of fluent linguistic production and a key factor in successful language learning. In other terms, these multi-word expressions include words that follow each other more frequently than expected by chance which shape the text meanings and are distinctive to a particular register. So, gaining control of a new language or register requires that expert users prefer certain sequences of words over others. Then, if learning to use more fixed phrases or lexical bundles of a discipline can contribute to gaining communicative competence in a field of study, there are advantages to identifying these clusters or bundles to better help

learners acquire specific rhetorical practices of their communities. Analysis of specialist corpora can therefore help us to understand the kinds of language data which particular communities of users might encounter and which will inform their use.

Understanding the lexical bundles perhaps sometimes is problematic for students of English language (literature, translation and teaching) because these can be a signsof ability of clear and strong writings and sometimes they are a writer`s outstanding style of writing therefore there are so many needs to be familiar with them and understand them clearly and fully. There is yet a lot to acquire about lexical bundles which might assist to discover an essential and nearly hidden aspects of genre analysis. Another dimension of the argument, which is not obvious yet is that how article authors` overuse, underuse, or adequate utilization of lexical bundles.

Although, there are so many studies and researches (as indicated in literature review) on lexical bundles, the researcher of this MA thesis tries to compare and contrast the differences and similarities of these usages just in the discussion sections of the ELT research articles focusing on the potential differences native and nonnative writers, an issue which has not been fully researched before. Two important characteristics about lexical bundles are their great frequency and their part in context building. Writing and speaking are never in vacuum it means when the writers and speakers try to have addresse they write for understanding and presenting their ideas to the others so it is logical they have their own clear and special style of writing. The study of native and non-native authors` use of lexical bundles can help students of the language and researchers, who want to use the results of the written papers, to be more familiar with the scientific texts that are presented in the articles and specially in the parts of discussions and results.

Literature Review

The definition of formulaic sequence from Wray (2006) is utilized broadly as reference in majority of studies. Formulaic sequence can be considered as a structure which might be both constant and inconsistent, and might be both of vocabularies or other factors, and that is, produced which means that is sustained and reobtained entirely from the recall for the time of utilization, rather than by examination by the language grammar.

Nevertheless, this meaning is topic to some criticism (Alquraishi 2014). The vagueness in the description results in the importance of human practice in realizing the formulaic sequence instead of machines, particularly if the formulaic sequence is inconsistent (Alquraishi 2014).

Similar to distinct vocabularies, lexical bundles differ through genres (Biber 2006), disciplines (Hyland 2008a) and authors' L1 (Lu & Deng 2019; Shin 2019). Their difference is of more importance in EAP setting which is presumed on the notion that any specific area of academic discourse is related to, or even limited to, specific linguistic characteristics, and effective EAP programs are the ones intended according to the particular requirements and discourse experiences of the target sets.

There are few indications that numerous lexical bundles are sustained as universal chunks and appreciate processing benefit above non-formulaic sequences (Nekrasova 2009). In a research that used self-paced reading activity, in the words of Tremblay et al. (2011) lexical bundles and statements encompassing lexical bundles were processed quick by the subjects than the control statement components, confirming that formulaic sequences are kept as distinct processing components. In another research, Jeong & Jiang (2019) utilized a subtle design and mentioned that only the structurally thorough lexical bundles were more rapidly processed. They made conclusion that for quicker processing, a lexical bundle must be not only extremely common but also structurally thorough.

Lexical bundles commonly appear in a register. Commonly, Cortes (2015) mentioned it is the final quality of lexical bundles. Nevertheless, making frequency onsets is somewhat random, and they varied from one research to another, mainly relying to corpus scope and manner of language.

So far a number of studies have been carried out in the field of lexical bundles. In this section a glimpse is taken to some of the recent ones.

Wachidah et al. (2020) attempted to examine the lexical bundles utilized in outcomes and discussion parts of graduate learners' thesis regarding structures, functions, and the part of lexical bundles in making coherence of the learners`writings. Their research utilized qualitative view. Structural classification of lexical bundles suggested by Biber et al. (1999), and the functions of lexical bundles suggested by Hyland (2008) outline were used. The results discovered that first, the lexical bundles discovered in the

learners` writings make utilization of all the twelve structural syntax and the most frequently utilized is structure kind 4 other prepositional phrase (component); second, the lexical bundles discovered in the learners` writings help all of the three functional kinds and the most frequently utilized is text-centered function and third, the lexical bundles discovered in the learners` writings have an essential role in creating coherence of the writings. It showed that the lexical bundles created utilization of two coherence elements like reference and transition indications.

Gil and Caro (2019) studied the utilization of lexical bundles in a learner corpus of undergraduate dissertations produced in English by Spanish L1 learners in linguistics and medicine, and contrasted it with issued papers in the similar fields. By concentrating on the introduction and conclusion parts, they recognized the most common 3-, 4- and 5-word bundles in the corpora, to later examine their kinds, structures, and functions. The outcomes indicated distinctions in the utilization of lexical bundles through fields, genres and parts, proposing pedagogical associations for the involvement of lexical bundles in the L2 composition curriculum.

Jalali and Moini (2018) worked on a research which considered a corpus-centered view for the realization of lexical bundles. A corpus of 801,894 vocabularies from 790 papers was gathered. In order to reach the objectives of the current study, ABBYY Fine Reader 10 professional edition, Total Assistant, Antconc 3.2.3, and WordSmith Tools 5 were utilized to recognize lexical bundles. The outcomes of the research showed that the authors of medical papers mainly depend on text-based bundles in the discussion part of papers to make academic discourse.

Jalilifar and Ghoreishi (2018) discovered two stages of formulaic sequences, overall and discipline particular, in a corpus of 200 applied linguistics papers and investigated the functions they imply. Utilizing Antconc software, in general, 2563 sequences were recognized containing 593 overall and 1370 discipline-particular classifications. Functional examination showed that overall and discipline-particular formulaic sequences monitor the similar design, with text-based sequences creating it to the highest part of the list tracked by research-based and participant-based sequences. Discipline-particular sequences exceeded overall sequences, presenting sign for the presence of more differences of discipline-particular formulaic sequences.

Hong and Hua (2018) recognized lexical bundles in the material of journal papers in the domain of international business management (IBM) and discovered that lexical bundles are field particular.

Gungor and Uysal (2016) contrasted the structure and function of the lexical-bundle utilization in L1 and L2 papers in English. The outcomes indicated that the standards of lexical bundles implications were varied between the non-native speakers of English and the native speaker standards. Moreover, the research indicated that Turkish investigators overworked clausal or verb-phrase concentrated lexical bundles when noun and prepositional phrase-centered lexical bundles were utilized more than clausal bundles by their native equivalents.

According to the background presented so far, and based on the existing gap in the literature the following research questions were proposed:

- RQ1: What are the most frequent four-word clusters used by native authors in ELT articles discussion section?
- RQ2. What are the most frequent four-word clusters used by non-native authors in ELT articles discussion section?
- RQ3. Is there any significant difference between the ELT research articles written by native and non-native authors, in terms of frequencies, structures, and functions of four-word clusters in their discussion sections?

Methodology

The present research was a corpus–based, descriptive study.

Corpus of the study

The corpus of this study was adopted from online journals for ELT professionals with open accessibility (ELT journal; TESL-EJ; L2 Journal; CALICO Journal; Foreign Language Annals; language teaching research; language teacher; ELT (Oxford); modern language journal; language learning; system; the language learning journal). All the ELT articles contained in this corpus were published during 2010 to 2020. Consequently, for the current research a written specific corpus including 200 written texts of a single genre, ELT articles with different authors (natives and nonnatives) was employed. In order to respond the questions of this research, discovering 4-word lexical bundles and their functions and

structures in discussion part of ELT articles, each part was given to the computer software to recognize lexical bundles in the articles. A summary of ELT journals was used in the study is presented in Table 1.

Table 1. *Description of the Corpus*

No	Journal	Number of Articles	Word Count
1	ELT journal	15	3208
2	TESL-EJ	16	3208
3	L2 Journal	17	3208
4	CALICO Journal	17	3208
5	Foreign Language Annals	15	3208
6	Language Teaching Research	14	3208
7	Language Teacher	17	3208
8	ELT (Oxford)	16	3208
9	Modern Language Journal	15	3208
10	Language Learning	15	3208
11	System	16	3208
12	The Language Learning Journal	14	3208

Instruments and used frameworks

Structural taxonomy of lexical bundles proposed by Biber et al. (1999). For structural categorization of 4-word lexical bundles, this study employed the framework proposed by Biber et al. (1999). Table 2 presents this categorization.

Table 2. *Structural taxonomy of lexical bundles by Biber et al. (1999: 1014)*

Category	Sample
Nounphrase +of	the end of the, the nature of the, the beginning of the, a large number of
Nounphrase +otherpost-modifier fragment	the fact that the, one of the most, the extent to which
Nounphrase+ prepositional phrase fragment	
Prepositionalphrase +of	at the end of, as a result of, on the basis of, in the context of
Other prepositional phrase	on the other hand, at the same time, in the present study, with respect to the
Anticipatoryit+ verb/adjective phrase	
Anticipatoryit+ adjective phrase	it is important to, it is possible that, it was found that, it should be noted
Anticipatoryit+ verb phrase	
Copulabe+ noun/adjective phrase	
Copulabe+ noun phrase	
Copulabe+ adjective phrase	
Verbphrase+ that-clause fragment	is the same as, is a matter of, is due to the, be the result of
Verbphrase+ that-clause	
Noun+ verb phrase+ that-clause	
Verb/adjective+ to-clause fragment	
Predicative adjective+ to-clause	as shown in figure, should be noted that, is likely to be, as well as the
Passive verb phrase+ to-clause	
To-clause	
Passiveverb+ prepositional phrase fragment	is shown in figure, is based on the, is defined as the, can be found in
Pronoun/nounphrase +be+….	
This+ be+…	
Otherexpressions	
Total	

Functional taxonomy of lexical bundles proposed by Hyland (2008)

The functional taxonomy of lexical bundles which was used in the present study was the taxonomy proposed by Hyland (2008) the details of which are presented in Table 3.

Table 3. *Functional taxonomy by Hyland (2008)*

Majorfunctions	Sub-categories	Examples
Research oriented: Help writers to structure theiractivities andexperiences oftherealworld	Location- indicatingtimeandplace	Inthepresentstudy,attheendof
	Procedure- Indicating methodology or purpose of research	The purpose of this, was used as
	Quantification- Describing the amount or number of	Is one of the, one of the most
	Description- Detailing qualities or properties of materials	In the control group, the size of the
	Topic related to the field of research	In the united states
Text-oriented: These clusters are concerned with the organization of the textand the meaning of its elements as a message or argument	Transition signals- establishing additive or contrastive links between elements	On the other hand, as well as the
	Regulative signals- mark inferential or causative relations between elements	The results of, been shown to be
	Structuring signals- text reflexive markers which organize stretches of discourse or direct reader elsewhere in text	As shown in fig, are shown in table
	Framing signals- situate arguments by specifying limiting conditions	On the other hand, in the presence of
Participant-oriented: These are focused on writer or reader of the text	Stance features-convey the writer's attitude and evaluations	Were more likely, it is possible that
	Engagement features- address readers directly	It should be noted, is important to note

Three main categories which are based on linguistic macro functions proposed by Halliday (1994) are: Research or real-world bundles concerned with ideational function, text-oriented bundles serve textual functions and participant-oriented bundles which are contributed to interpersonal functions.

Procedure

As it was mentioned earlier, the main aim of this research was to recognize 4-word lexical bundles in the corpus of discussion part of ELT articles with different authors (natives and nonnatives) and to investigate the recognized lexical bundles structurally and functionally. Consequently, two major sections were conducted in this study. Based on the aim of current research, in the first part, all lexical bundles were identified and were classified structurally based on Biber et al. (1999) taxonomy and functionally using Hyland (2008a, 2008b) functional category of lexical bundles in academic text. In the second section, all main bundles recognized in discussion part of ELT articles that detected two standards of frequency (occurrence of 20 times per million word) and distribution (occur in at least five different texts) were grouped structurally and functionally utilizing the above mentioned categories respectively. After that, a contrastive study was carried out among main bundles considering the functions they work utilizing *Chi square* means to realize if frequency diversions in utilizing lexical bundles are statistically substantial through the corpora.

Antconc software (version 3.3.2), proposed by Anthony (2011), was used in order to analyze the corpora. To be more exact, the following stages were done to carry out this research. First, an exclusive code was given to every discussion part (e.g., A#1, A#2, A#3… A#200). Second, the corpora classified into two category of discussions (i.e., each category contained of 100 discussions). The two categories were given to the program one by one and were investigated structurally according to Biber et al. (1999) classification and functionally utilizing Hyland (2008a, 2008b) functional category of lexical bundles in academic text. In order to check the reliability of the gathered data, a PhD candidate was asked to score 30% of the data, and a Cronbach alpha indicated a reliability index of 0.94.

Results

Frequency of Bundles in Native ELT Research Articles

Table 4 presents the frequency of four word lexical bundles in the corpus of ELT research articles authored by native researchers.

Table 4. *Most Frequent Four-word Bundles in Native Articles*

	Lexical bundles	F	p
1	the results of the	9	6.56
2	on the other hand	13	9.48
3	in the case of	6	4.37
4	in the target language	7	5.1
5	the meaning of the	2	1.45
6	at the same time	8	5.83
7	in the process of	2	1.45
8	at the end of	1	.72
9	is one of the	1	.72
10	as well as the	7	5.10
11	on the basis of	9	6.56
12	the results of this	10	7.29
13	in the present study	9	6.56
14	in other words the	3	2.18
15	in the use of	2	1.45
16	results of this study	9	6.56
17	to the fact that	7	5.10
18	the end of the	5	3.64
19	as a result of	5	3.64
20	as a foreign language	2	1.45
21	the findings of the	3	2.18
22	used in this study	10	7.29
23	the analysis of the	7	5.10
	Total	137	

As shown, a total number of 137 bundles were found in the corpus of native ELT articles which were fewer than the bundles found in the nonnative corpus. The most highly frequent lexical bundle in the native ELT research

articles discussion section was the bundle of *on the other hand* with a frequency of 13 and a percentage 9.48%. An example of this lexical bundle is presented here.

The method factor loadings, on the other hand, are largest for the reading translation (0.879) and for the self-ratings (0.437 and 0.519).

Two other bundles, namely, used in this study and the results of this were employed identically in the corpus of native ELT articles both with a frequency of 10 and a percentage of 7.29%. Instances of these two bundles are presented below:

Finally, with respect to analyses employed, the factor analyses (both exploratory and confirmatory) used in this study are ….

…., the results of this study show that the prediction was confirmed.

The bundles of the results of the, on the basis of, in the present study and results of this study were used identically with a frequency of 9 and a percentage of 6.56%. The following sentences include instances of these bundles:

The results of the study do point strongly to certain tendencies among teachers.

On the basis of a scree test, three factors were identified.

In the present study, we derived scores for four empirically and theoretically based factors consistent with Moore et al. (2015) using CFA.

The next frequent four-word lexical bundle was *at the same time* with a frequency of 8 and a percentage of 5.83%. An example of this bundle is presented in the following sentence.

At the same time, they were able to attend to at least some of the target words, take note of their form, and use the context to help them establish a form-meaning mapping.

The next four frequent lexical bundles were in the target language, to the fact that, as well as the, and the analysis of the with frequency of 7 and percentage of 5.1%.

Another lexical bundle was found to be *in the case of* with a frequency of 6 and a percentage of 4.37%. The following sentence include an example of this lexical bundle.

In the case of other sorts of feedback, including recasts, interaction researchers have claimed that learners might sometimes repeat the feedback they received.

In addition, the bundles of the end of the and as a result of were the next frequent bundles with frequencies of 5 and percentage of 3.64%.

Finally, the rest of four word bundles were used less than 5 times in the corpus of native ELT research articles. Some examples of the employed lexical bundles are presented in the following sentences:

The more automatic their access, the more affluent is the resultant language use, concomitantly freeing attentional resources for analysis of the meaning of the message, either for comprehension or for production planning. Figure 1 presents the five most frequent bundles in the corpus of native articles.
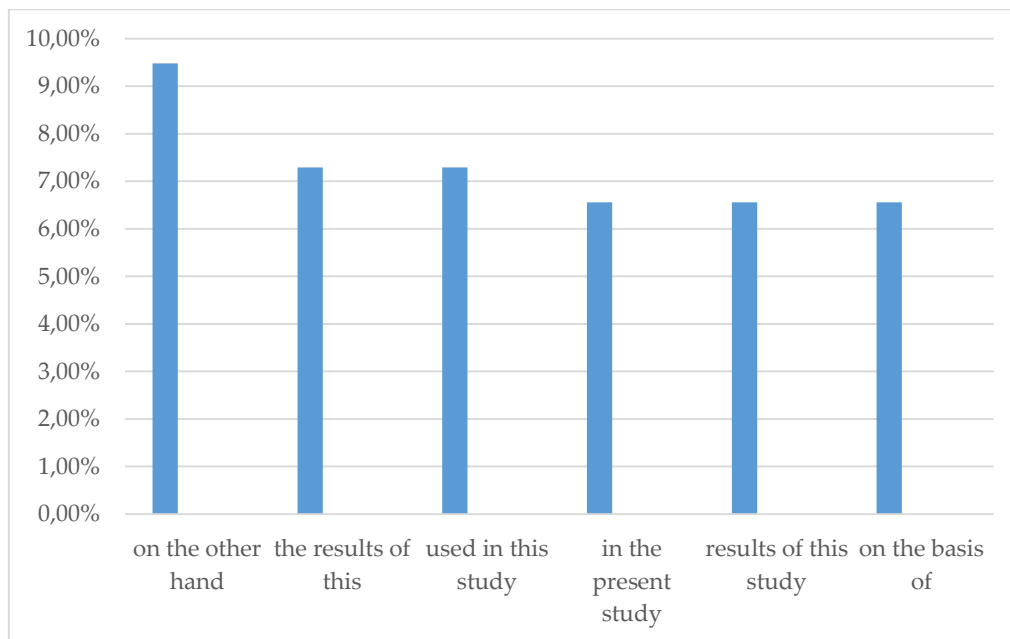

Figure 1. Most Frequent Four-word Bundles in Native Articles

4.2 Frequency of Bundles in Non-native ELT Research Articles

The result of analysis of four word bundles in the nonnative corpus is presented in Table 5.

Table 5. *Most Frequent Four-word Bundles in Non-Native Articles*

	Lexical bundles	F	P
1	on the other hand	21	10.44
2	in the case of	5	2.48
3	the results of the	41	20.39
4	as well as the	5	2.48
5	in the context of	4	1.99
6	it is important to	3	1.49
7	at the end of	5	2.48
8	the end of the	7	3.48
9	as a result of	14	6.96
10	can be found in	3	1.49
11	that there is a	9	4.47
12	the use of the	3	1.49
13	should be noted that	10	4.97
14	can be seen in	5	2.48
15	in terms of the	7	3.48
16	in the form of	11	5.47
17	it should be noted	11	5.47
18	the total number of	3	1.49
19	at the same time	13	6.46
20	in the present study	15	7.46
21	that there is no	6	2.98
	Total	201	100

The four word lexical bundles in the non-native corpus are also presented in Table 5. As shown the total number of four word bundles were 201. According to the results, the most highly frequent bundle was the bundle of *the results of the* with a frequency of 41 and a percentage of 20.39%. An example of this bundle in the corpus is presented in the following sentence.

This confirms the results of the quantitative phase and also the findings of other L2 DA studies like Shrestha and Coffin (2012), and Poehner (2009).

The next frequent bundle was that of *on the other hand* with a frequency of 21 and a percentage of 10.44%. The following sentence includes an instance of this bundle.

In Australia, on the other hand, applicants are able to study and teach at the same time but the studying hours are set as they can easily manage their time for both practice and studying.

In the present study was the next frequent bundle with a frequency of 15 and a percentage of 7.46%. An example of this bundle in nonnative ELT corpus is observed in the following sentence.

The experts participated in the present study emphasize that the lack of condition in stage completion or success rate of completion might be misleading for the language learners.

As the next frequent bundle, As a result of accounted for 6.96% of the whole nonnative corpus. An instance of this bundle is presented in the following sentence.

Similarly, the results seem inconsistent with the findings of the studies by Carrell and Wise (1998) who reported no significant impact of topic interest on the learners\x92 L2 reading comprehension, whereas the results of the present study showed an improvement in learning as a result of using the interest-based teaching.

The next four-word lexical bundle in nonnative corpus was at the same time with a frequency of 13.

In Australia, on the other hand, applicants are able to study and teach at the same time but the studying hours are set as they can easily manage their time for both practice and studying.

In addition, in the form of and it should be noted were found identically with frequencies of 11 and percentages of 5.47%.

The next frequent bundle was should be noted that which accounted for 4.97% of the whole corpus. An instance of this bundle is shown in the following sentence.

It should be noted that the assumption of homogeneity of variances was met (Levene F = 2, p = .161 > .05).

The rest of the four word lexical bundles were used less than 5%. In order to have a more comprehensive picture about the issue under study, Figure 2 presents the five most frequent bundles in the corpus of nonnative articles discussion section.
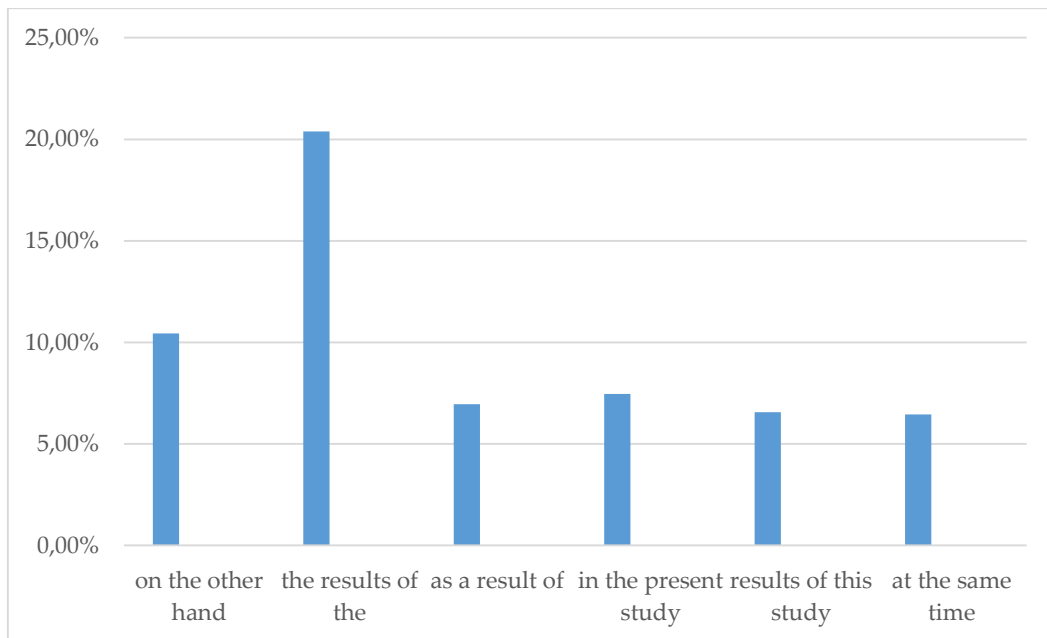

Figure 2. Most Frequent Four-word Bundles in Non Native Articles

In order to make sure of the (in) significance of the difference between the lexical bundles of the two groups, a Chi-square test was administered among the common lexical bundles of the two groups. Table 6 presents the results.

Table 6. *Results of Chi-square Test between Four-word Bundles in Native and Non Native Articles*

	Value	DF	Asymptotic Significance (2-sided)
Pearson Chi-Square	73.746[a]	8	.000
Likelihood Ratio	1.815	6	.936
Linear-by-Linear Association	.079	1	.779
N of Valid Cases	1362		

According to the results presented in Table 6, there was a significant difference between the four word lexical bundles in the corpus of native and nonnative ELT articles discussion sections in terms of their frequency ($p<.05$).

Structural classification of bundles in native and non-native ELT research articles

Table 7 shows the structural classification of 338 identified lexical bundles in the corpus of discussion sections of ELT research articles. The frequency of occurrence of each bundles is presented.

Table 7.

StructuralClassification of Lexical Bundles in the Conclusion Section of ELT Research Articles

Category	Native		Nonnative	
	F	p	F	p
Nounphrase + of	20	12.9	25	13.66
Nounphrase +other post-modifier fragment	5	3.22	6	4.34
Nounphrase+ prepositional phrase fragment	5	3.22	6	4.34
Prepositionalphrase +of	5	3.22	5	2.73
Other prepositional phrase	5	3.22	6	4.34
Anticipatoryit+ verb/adjectivephrase	10	6.45	11	7.97
Anticipatoryit+ adjectivephrase	5	3.22	6	4.34
Anticipatoryit+ verbphrase	5	3.22	5	2.73

Category	Native		Nonnative	
	F	p	F	p
Copulabe+ noun/adjective phrase	20	12.9	25	13.66
Copulabe+ nounphrase	10	6.45	10	5.46
Copulabe+ adjective phrase	10	6.45	15	8.19
Verbphrase+ that-clause fragment	20	12.9	25	13.66
Verbphrase+ that-clause	10	6.45	10	5.46
Noun+ verb phrase+ that-clause	10	6.45	15	8.19
Verb/adjective+ to-clausefragment	30	19.35	20	10.92
Predicative adjective + to-clause	10	6.45	7	3.83
Passive verb phrase + to-clause	10	6.45	7	3.83
To-clause	10	6.45	6	4.34
Passiveverb + prepositional phrase fragment	10	6.45	20	10.92
Pronoun/nounphrase +be+….	10	6.45	11	7.97
This+ be+…	10	6.45	11	7.97
Otherexpressions	25	16.12	29	15.83
Total	155		183	100

Note: Those categories thatare in bold are main categories, and the others are subcategories

As shown in Table 7, a total number of 338 lexical bundles were identified in the corpus under study including a number of 155 bundles used by native and 183 bundles used by nonnative bundles. The lexical bundles were categorized into 11 categories which were used by both native and nonnative writers. From among the categories Verb/adjective+ to-clause fragment was used more frequently than other categories by native writers with a percentage of 19.35%. This category was employed by nonnative

writers with a percentage 10.92%. This category involves Predicative adjective+ to-clause, Passive verb phrase+ to-clause and to-clause was used with percentages of 6.45%, respectively. The category of Noun+ verb phrase+ that-clause was used with a percentage of 6.45% by native writers and 8.19% nonnative writers. An example of this class of lexical bundle is presented here: This paper showed that the, and these findings support that.

The next class of lexical bundles was verb phrase + that clause fragment with a percentage of 12.9% by native writers and 13.66% by nonnative writers, which included three sub-categories, namely, verb phrase + that clause and Noun+ verb phrase +that clause. Verb phrase + that clause. Examples of this bundle can be seen in the following samples haverevealedthatthe, should bepointed out thattheandshouldbe noted that. Another category was Nounphrase + of with a percentage of 12.9% by native writers and 13.66% by nonnative writers. The next category of lexical bundles found in discussion sections of ELT research articles was the category of Nounphrase +otherpost-modifier fragment with 3.22% in the native corpus and 4.34 by nonnative writers.

The next most frequent type among clausal bundles is related to thecategory of copula be + noun/adjective phrase which constitutes 12.9% of thewhole bundles used by native authors and 13.66% of the whole bundles by nonnative writers. Examples of this category were is possible that the, is important to note, isalsopossiblethat.

The next highly frequent category of lexical bundle was Anticipatory it+ verb/adjective phrase with a percentage of 6.54% for native corpus and 7.97% for nonnative corpus.

Passive verb+ prepositional phrase fragment formed the next category of lexical bundles with a percentage of 6.45% in the native and a percentage of 10.92% in the nonnative corpus. Figure 3 presents the results.
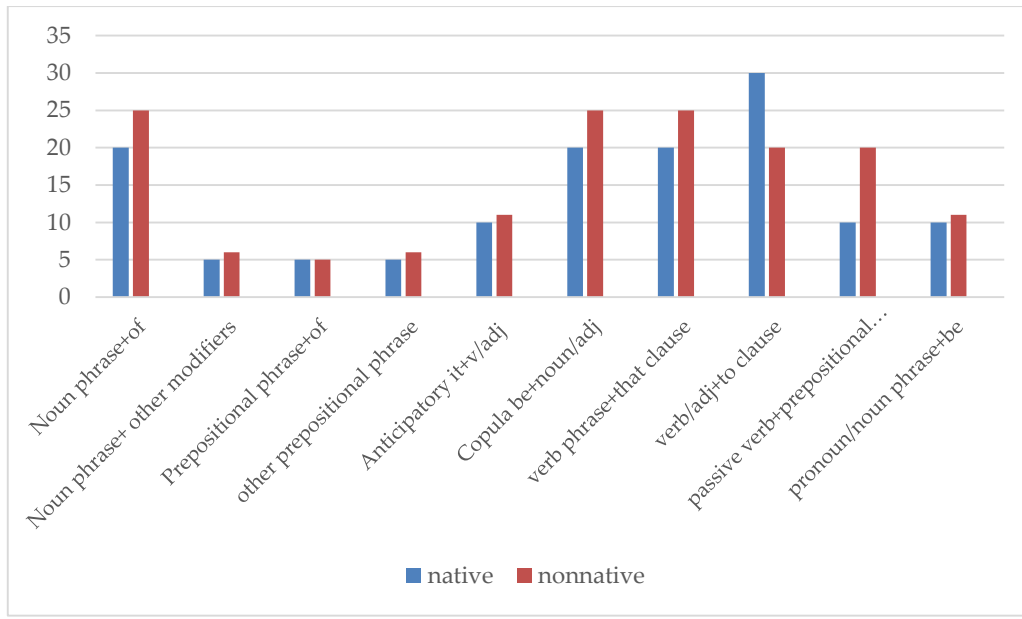
Figure 3. Structural Classification of Lexical Bundles in the Discussion Sections

In order to investigate if there were any significant differences in structural classification of lexical bundles in discussion sections of ELT research articles, the gathered data were exposed to Chi-square test. The results are presented in Table 8.

Table 8. *Results of Chi-Square Test in Structural Classification of Lexical Bundles*

	Value	DF	Asymptotic Significance (2-sided)
Pearson Chi-Square	27.32	9	.031
Likelihood Ratio	23.02	9	.031
Linear-by-Linear Association	3.23	1	.01
N of Valid Cases	1123		

Table 8 shows the results of the Chi-square test. As shown, the level of significance is smaller than the alpha level $P > 0.05$; accordingly, there were significant differences among the lexical bundles in the discussion sections of ELT research articles written by native and non-native scholars in terms of their structural classifications.

Functional Analysis of Bundles in Discussion Sections of ELT Articles

Hyland's taxonomy (2008a) was employed to analyze the gathered data functionally. In total, functional classification of 338 bundles identified in the corpora of ELT research articlesdiscussion sections were presented. The results are presented in Table 9.

Table 9. *Functional Classification of Bundles in ELT Articles*

Categories	Subcategories	Native		Non native		Samples
		F	p	F	p	
	Total	23	6.8%	47	13.9	
Research-oriented bundles	Location	2	.59	4	1.18	in the context of
	Quantification	4	1.18	8	2.36	the total number of
	Procedure	3	.88	9	2.66	the use of the
	Study- focusing	9	2.66	17	5.02	In the present study
	Evaluation	3	.88	5	1.47	The validity of the
	Discipline-bundle	2	.59	4	1.18	In the context of
Text-oriented bundles	total	68	20.11	122	36.09	
	Resultative-Signals	15	4.43	29	8.57	can be found in, the result of the
	Framing –signals	8	2.36	19	5.62	On the other hand
	Reference	9	2.66	18	5.32	mentioned in previous studies
	Confirmation	15	4.43	23	6.8	This is in line with
	Suggestion	13	3.84	20	5.91	Studies are needed to
	Relation	8	2.39	13	3.84	Be related to the
Participant-oriented bundles	Total	27	7.98	51	15.08	
	Attitude markers	5	1.47	9	2.66	It is reasonable to
	Epistemic-certain	5	1.47	9	2.66	It is clear that
	Epistemic-Uncertain	9	2.66	21	6.25	It is likely that
	Engagement features	8	2.36	12	3.57	It is important to
Total		118	34.91	220	65.09	

According to the descriptive statistics, text-oriented bundles were found to be the main function for which lexical bundles were used in ELT research articles by both native and nonnative scholars. In fact, of the whole number of bundles in the corpora under study, 20.11% were allocated to this function by native writers and 36.09% by nonnative writers. The findings

showed that that text-oriented lexical bundles in ELT research articles were composed of six types of functional categories. Resultative signals were found to be the main function in the category of text oriented bundles in comparison with other subcategories. This subcategory of lexical bundle function accounts for 36.09% of the whole number of lexical bundles used by non-native writers and 20.11% of the bundles used in ELT research articles written by native writers.

Confirmation bundles which are used by writers to enhance the reliability or validity and correctness of results was found to be the next category of lexical bundles with a frequency of 23 among ELT articles by non-native writers and a frequency of 15 among those articles written by native authors. Examples of this category of bundles can be seen in expressions such as this is in line with.

Another sub-category used by both native and non-native authors was found to be suggestion bundles. As the name of this type of bundles suggests, the lexical bundles in this functional category are mainly used by authors of to refer to the suggestions for further researches. The findings showed that 20 bundles (5.91% of the text oriented lexical bundles employed by nonnative writers) and13 bundles (3.84% of the bundles used by native writers) were included in this functional category. A sample of this functional category in the corpora was studies are needed to.

As the next functional categories of lexical bundles, identically Framing signals and References were found both with percentages of 5.62% and 5.32% for nonnative writers and 2.36 % and 2.66% for native writers, respectively.

The last sub-category of text-oriented bundles was found to be relation bundles. This class of bundles are used to showthe relation between two or more elements in a study. This bundle was used with an overall percentage of 3.84% by nonnative writers and 2.39% by native writers. An example of thesebundles in the corpora was the phrasementionedinprevious studies.

In addition, as shown in Table 4.1, research-oriented bundles constitute 20.07% ofbundles in ELT research articles. In this category of lexical bundles, study-focusing bundles formed thehighest frequency of bundles with a percentage of 5.02% for nonnative writers and 2.66% for native writers. Study-focusing bundles referexplicitly to the study being

conducted and reported by writers of ELT articles among which in the present study was considered as the most frequent bundle with the highest frequency of 17 by nonnative writers and 9 for native writers.

The lexical bundles category of location represent the least frequently used category intext-oriented bundles. In the context of is an example of this functional subcategory.

Evaluation bundles which refer to some of the evaluations that are made bythe researcher during the study. This bundle with an overall percentage of 1.47% of the wholebundlesinthecorpora of non-native writers, and a percentage of .88% of the bundles by native writers. Theleastsub-category in research orientedbundles is related to discipline-boundbundles that are distinctive common word combinations in a specific field.

Participant-oriented bundles comprised 23.06% of the whole bundlesamong which the highest percentage is recommended by epistemic-uncertain features and attitude markers which accounted for 2.66% of whole number of lexical bundles for nonnative and 1.47% for native writers. An example of this function was notableto and it isclearthat.

Epistemic-certainbundles, which express a more certain stance toward their propositions was used with a percentage of 6.25% of the whole number of lexical bundles in the corpora of nonnative ELT articles and 2.65% in the corpus of native articles. Examples of these two bundles includeitisclearthatandit is likely that respectively.

Engagement bundles which are classified under the category of participant-oriented bundles are used to address the reader directly. Finally, attitude-markers were employed with a frequency of 1.47% of the whole bundles in the corpora of native and non-native articles. In order to be more objective about the functional analysis of lexical bundles, Figure 4 compares the functions of the lexical bundles used in the ELT research articles.
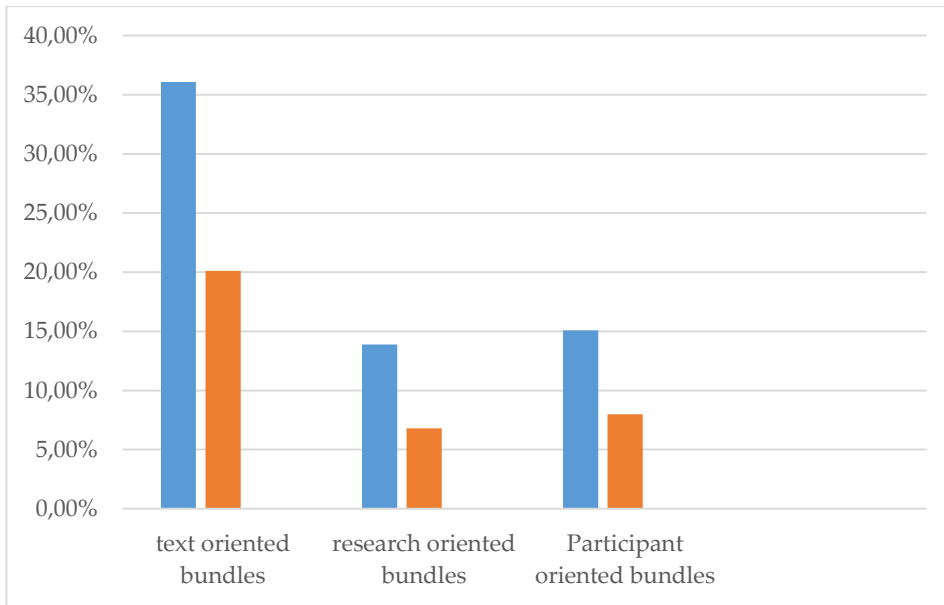
Figure 4. Functional Classification of Four Word Bundles in ELT Articles

As depicted in Figure 4, text orientation was identified as the most significant function used by native and non-native writers, for which four word lexical bundles were adopted. The next functional category of bundles was the Participant orientation, and the last function of lexical bundles was participant orientation in ELT research articles. In order to compare native and non-native writers in terms of these functions a *Chi-square* test was administered. Table 10 presents the results.

Table 10. *Results of Chi-SquareTest between Functional Classifications of Bundles in ELT Articles*

	Value	DF	Asymptotic Significance (2-sided)
Pearson Chi-Square	24.52	2	.00
Likelihood Ratio	24.65	2	.00
Linear-by-Linear Association	5.77	1	.01
N of Valid Cases	1303		

Table 10 shows the results of the *Chi-square* test. As shown, the level of significance is smaller than the alpha level P > 0.05; accordingly, there were significant differences among ELT articles written by native and non-native writers in terms of their functions.

Discussion

The results revealed that authors of both corpora made use of various types of four-word lexical bundles; however, there were significant differences between native and nonnative articles in terms of their use of four-word lexical bundles. In fact, Iranian authors used theses lexical bundles almost twice more than native authors. This can be justified in light of the settings in which the articles were written. Due to the effect of first language, native writers do not need to overuse different lexical bundles, while nonnative writers need to use lexical bundles with more frequency to get their message crossed.

The results of different frequencies of functional categories in native students' writing lend support to Hyland's study (2008a). One more finding of the present thesis was that nonnative writers made use of research oriented bundles more than native writers. This can be compared with the study by Jalali (2009). In which it was found that Iranian students use research bundles more than other categories. And participant-oriented clusters are the least frequent ones. Furthermore, in the present thesis it was found that participant oriented bundles were used less than other bundles. This may be due to the fact that academic articles are usually written by scholars with high level of knowledge who usually bring their ideas into their research by using participant oriented bundles. This is in line with the study by Jalali (2008).

In terms of functions of four-word lexical bundles, it was found that text-oriented bundles were the main function for which both native and nonnative authors of ELT articles use lexical bundles. In this regard, it can be stated that the discussion section of articles deals with the organization of the text through which the researchers can send a message or set an argument; therefore, they make use of text oriented bundles more than other bundles. The findings in this regard lend support to the study by Hyland (2008), in that in both studies, the mostly used functional category in both groups is 'text oriented bundles'. Considering text oriented bundles. In this regard, the findings of the present thesis lend support to the study by Hyland (2008a) in which it was found that in native students' writings the mostly used functional category was text oriented bundles. The reason for such findings can be explained in light of the fact that authors are usually more interested in talking about the methodology they have

employed in their writing. This finding are contrary to the findings of the study by Jalali (2009). This may be due to the fact that the participants of the present thesis were different from the study by Jalali (2009).

In addition, the results of the present thesis in this regard lend support to the study by Jalali & Moini (2018) in which they investigated lexical bundles in the discussion section of Medical Research articles and showed that medical authors depend on text-oriented bundles in the discussion section of research article to establish academic discourse.

In addition, participant oriented bundles such as that epistemic-certain bundles were used by both native and nonnative authors. Biber & Barbieri (2007) stated that *epistemic-certain* bundles form aframe and project the propositions as unhedged and undisputed arguments. Differently stated, *epistemic-certain* bundles are about the certainty of the writer toward what he/she istalkingabout. In fact, most authors use epistemic certain bundles to avoid making a mistake and being accused by others (Jalali 2009). Moreover, both groups of authors employed engagement markers. This finding can be justified in light of the fact that writersusethis type of engagement marker to engagethereaderwith thetext andattract her/his attention to a particular point. Attitude markers were also found in both corpora. Attitude markers are used to show writers' overtstance toward a subject or what she/he is talking about.

In terms of structural classification of bundles, it was found that prepositional phrase + of" was used less than other bundles. This was comparable with the study Amirian (2013); however, it was in contrast with the study by Hyland (2008). In Hyland's study, the most frequent clusters were found to be other prepositional phrases. In the present study in general it was found that phrasal bundles were more common than other types of bundles. This is in line with other studies such as Biber et al. (1999). The reason for higher number of phrasal bundles in comparison with clausal bundles might be because of vague structural classification of lexical bundles.

Conclusion

According to the results, it can be concluded that lexical bundles are considered as inseparable constituents of academic discourse. The high frequency of lexical bundles in the corpus used in this study proves such

claim. Probably, many researchers whose research articles were analyzed in literature review, had no difficulty in understanding lexical bundles considering the fact that they may have been encountered with various types of clusters quite often in their readings earlier; however, the authors may have tried to avoid some of them on the basis of a mistaken assumption that the use of such word sequences may signal unconfirmed claims, weak language or even impoliteness. There is almost no doubt that such researchers have repeatedly observed different lexical bundles in different research articles they may have studied for doing and writing their own research. Furthermore, given that lexical bundles are very prevalent in university written language and they may have a formulaic status, it may be expected the acquisition of such word combinations may not confront students with a very difficult task given their relatively high level of language proficiency and disciplinary writing expertise.

Implications of the Study

One of the implications of this study is that through the proper use of lexical bundles, the academia can improve their writing, and the role of these bundles in transferring the data and achieving the goal of teaching is of great significance. The study of native and non-native authors' use of lexical bundles can help students of the language and researchers, who want to use the results of the written papers, to be more familiar with the scientific texts that are presented in the articles and specially in the parts of discussions and results. The findings of this thesis can be of advantage to the instructors and learners of academic writing in that it suggests a wide range of resources for framing the language of research articles, and it assists managing the available resources in developing research articles. Hence, the results help us with knowledge construction on the ways in which various bundles are employed in the field of ELT research articles.

The findings of the present research can also have implications for academia in general and researchers in the fields of hard sciences in particular to become aware of different conventions governing the genre of academic papers. Various word clusters including lexicalbundles are used in divergent ways in different languages and disciplines; this makes it important for teachers and ESP as well as EAP course designers to recognize this for instruction. These findings reveal what linguistic and discourse features should be taught in the classrooms, and more certainly, what

should be incorporated into the ESP and EAP courses and textbooks. In fact, the findings can be used as guidelines to feed the courses of English for Academic/Specific Purposes in terms of syllabus and curriculum content and design.

Competing interests: The author(s) declare none.

References

Alquraishi, M. A. (2014). *Lexical bundles in an advanced INTOCSU writing class and engineering texts: A functional analysis*. Doctoral dissertation, Colorado State University.

Amirian, Z. (2013). The use of lexical bundles in native and non-native post-graduate writing: the case of applied linguistics MA theses. *Journal of English Language Teaching and Learning, 5*(11), 1-29.

Anthony, L. (2011). AntConc (Windows, Macintosh OS X, and Linux). Recuperado de: http://www. antlab. sci. waseda. ac. jp/software/README_AntConc3, 2.

Biber, D. (2006). *University Language: A Corpus-based Study of Spoken and Written Registers*. Unpublished, Amsterdam: John Benjamins.

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.

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

Biber, D., & Gray, B. (2010). Challenging stereotypes about academic writing: Complexity, elaboration, explicitness. *Journal of English for Academic Purposes, 9*, 2-20.

Biber, D., Gray, B., & Poonpon, K. (2011). Should we use characteristic of conversation to measure grammatical complexity in L2 writing development? *TESOL Quarterly, 45*(1), 5-35.

Biber, D., Conrad, S., & Cortes, V. (2004). If you look at ...: Lexical bundles in university teaching and textbooks. Applied *Linguistics, 25*(3), 371-405.

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *The Longman Grammar of Spoken and Written English*. London: Longman.

Carrell, P. L., Gajdusek, L. & Teresa, W. (1998). Metacognition and EFL/ESL reading. *Instructional Science*, 26(1), 97-112.

Cortes, V. (2015). Situating lexical bundles in the formulaic language spectrum: Origins and functional analysis development. In V. Cortes, & E. Csomay (Eds.), *Corpus-based research in applied linguistics: Studies in honor of Douglas* Biber (pp. 197e216). Amesterdam: John Benjamins.

Gil, N. N., & Caro, E. M. (2019). Lexical bundles in learner and expert academic writing. *Bellaterra Journal of Teaching & Learning Language & Literature*, *12*(1), 65-90.

Gungor, F., & Uysal, H. H. (2016). A comparative analysis of lexical bundles used by native and non-native scholars. *English Language Teaching*, *9*(6), 176-191.

Hong, A. L., & Hua, T. K., (2018). Specificity in English for Academic Purposes (EAP): A Corpus Analysis of Lexical Bundles in Academic Writing. *3L: The Southeast Asian Journal of English Language Studies, 24*(2),82 – 94.

Hyland, K. (2008a). As Can be Seen: Lexical Bundles and Disciplinary Variation. *English for Specific Purposes, 27*, 4-21.

Hyland, K. (2008b). Academic clusters: Text patterning in published and postgraduate writing. *International Journal of Applied Linguistics, 18*(1), 41-62.

Hyland, K. (2012). Bundles in academic discourse. *Annual Review of Applied Linguistics, 32*, 150-169.

Jalali, Z. S., & Moini, M. (2018). A Corpus-based Study of Lexical Bundles in Discussion Section of Medical Research Articles. *Iranian Journal of Applied Language Studies*, *10*(1), 95-124.

Jalilifar, A., & Ghoreishi, S. M. (2018). From the perspective of: Functional Analysis of Lexical Bundles in Applied Linguistics Research Articles. *International Journal of English Studies*, *18*(2), 161-186.

Jeong, H. & Jiang, N. (2019). Representation and processing of lexical bundles: Evidence from word monitoring. *System 80,* 188-198.

Lu, X., & Deng, J. (2019). With the rapid development: A contrastive analysis of lexical bundles in dissertation abstracts by Chinese and L1 English doctoral students. *Journal of English for Academic Purposes, 39*, 21-36.

Nekrasova, T. M. (2009). English L1 and L2 speakers' knowledge of lexical bundles. *Language Learning, 59*(3), 647-686.

Poehner, M., E. (2009). Group dynamic assessment: Mediation for the L2 classroom. *TESOl Quarterly, 43*(3), 471-491.

Shin, Y. K. (2019). Do native writers always have a head start over nonnative writers? The use of lexical bundles in college students' essays. *Journal of English for Academic Purposes, 40*, 1-14.

Shrestha, P. & Coffin, C. (2012). Dynamic assessment, tutor mediation and academic writing development. *Assessing Writing, 17*(1), 55-70.

Tremblay, A., Derwing, B., Libben, G. & Westbury, Ch. (2011). Processing advantages of lexical bundles: Evidence from self-paced reading and sentence recall tasks. *Language Learning 61*(2), 569-613.

Wachidah, W. D. N. A., Fitriati, S. W., &Widhiyanto, W. (2020). Structures and Functions of Lexical Bundles in Findings and Discussion Sections of Graduate Students' Thesis. *English Education Journal*, *10*(2), 131-142.

Ward, A. & Litman, D. J. (2007). *Automatically measuring lexical and acoustic/prosodic convergence in tutorial dialog corpora.* In: SLaTE Workshop on Speech and Language Technology in Education, 01 October 2007-03 October 2007, Farmington, PA.

Wray, A. (2006). Formulaic language. In E. K. Brown, & A. *Anderson (Eds.), Encyclopedia of language and linguistics* (pp. 590-597). Boston: Elsevier.