Journal of Critical Applied Linguistics Studies

ORGINAL RESEARCH
E-ISSN: 2981-1678
Volume 1, Issue 1, pp. 87-104

Gonbad Kavous University

Lexical Bundles in the Conclusion Section of Applied Linguistics Articles: A Comparative Study of International & Iranian Journals

Linguistics Society of Iran

[1]*Ghorbani Ali* [2]Farideh Okati* [3]Behruz Lotfi Gaskaree

ABSTRACT

Lexical bundles, sequences of words that recur frequently in corpus-based discourse, are indicators of fluent language production both in L1 and L2. Hence, to the best of the researcher's knowledge, no research has been conducted on the characteristics of lexical bundles in the conclusion section of applied linguistics research articles. Consequently, the present study aims to extend this line of research by comparing the form, function, and frequency of four-word lexical bundles in the conclusion section of 225 research articles published in international vis-à-vis Iranian applied linguistics journals. The findings revealed that the bundles were more frequent in the conclusion section of research articles in Iranian journals as opposed to international journals. Structural analysis of lexical bundles showed that passive structure was used infrequently in international journals. In addition, analyzing the function of lexical bundles indicated that procedure and resultative bundles were used more frequently by Iranian writers. Based on the results of this study, it can be concluded that Iranian writers overused lexical bundles in terms of passive structure, procedure, and resultative bundles in comparison with international writers. Results will be applicable for researchers and authors who want to communicate more effectively in their communities of practice.

Article History

Received:
2023-08-02
Revised:
2023-09-25
Accepted:
2023-10-27
Published:
2024-01-01

Key Words:
Applied linguistics, Conclusion section, formulaic language, International applied linguistics journals, Iranian applied linguistics journals

1. M.A in ELT, Department of English Language, University of Zabol, Zabol, Iran. *Email: ali.ghorbani@live.com,* ORCID iD: https://orcid.org/0000-0002-2027-2308

2. **Corresponding Author:** Assistant Professor of Linguistics, Department of English Language, University of Zabol, Zabol, Iran. Email: *farideh.okati@uoz.ac.ir,* ORCID iD: https://orcid.org/0000-0001-7399-8551

3. Assistant Professor of Linguistics, Department of English Language, University of Zabol, Zabol, Iran. *Email: lotfi@uoz.ac.ir,* ORCID iD: https://orcid.org/0000-0002-2413-3620

Article Citation: Ghorbani, A., Okati, F.*, & Gaskaree, B. L. (2024). Lexical bundles in the conclusion section of applied linguistics articles: A comparative study of international & Iranian journals. *Journal of Critical Applied Linguistics Studies*, *1*(1), 87-104. http://dx.doi.org/10.22034/jcals.1.1.67.

[DOI: 10.22034/jcals.1.1.67]

1. Introduction to the Problem

Corpus-based studies revealed that although native speakers produce multi-word units naturally, such units are one of the main problematic areas for non-native speakers (De Cock, 2003; Nesselhauf, 2005), which can also hinder the production of a language. Moreover, in achieving native-like competency and proficiency, these LBs are vital and are hence a significant point that needs to be brought into spotlight in educating and acquiring a language (Coxhead, 2008; O'Keeffe, McCarthy, & Carter, 2007). One of the principles of good writing is using the words in the correct context and combinations. Therefore, knowing the most occurring bundles used in specific disciplines, registers, and genres for second or foreign language writers is necessary. This is also very important in scientific writing, since producing a concise, exact, and precise writing is necessary for writers to be able to convey their thoughts and research results to their special audience.

Recently, genre studies have highlighted the term "lexical bundle", a type of multi-word expressions, as building blocks for spoken and written expression (Biber et al. 1999; Biber & Conrad, 1999; Biber, Conrad & Cortes, 2004; Cortes, 2004, 2006; Hyland, 2008a; Chen & Baker, 2010). Lexical bundles are clearly defined as the most frequently repeated word sequences (e.g., *I don't know if, I just wanted to*). It is discipline-bound in published academic writing (Cortes, Jones, & Stoller, 2002), which indicates that every discipline has its objective and ways of seeing the world related to specific conventions of a particular community (Hyland & Hamp-Lyons, 2002). *It can be concluded, in the present study,* and *the result of the,* among many others, are instances of lexical bundles commonly found in the academic literature. Scholars and learners who regularly use a certain genre are familiar with these terms, and thus, they signify being a competent engagement in a given user community. The absence of such clusters, on the other hand, may demonstrate that the participant of the target community is not fluent. Therefore, lexical bundles make a distinction between the use of written and spoken languages by novices and professionals in different contexts, allowing participants in a discourse community to express solidarity with other participants.

Over the past 20 years, researchers have paid considerable attention to LBs, and lots of research has been done in this area. Researchers have analyzed LBs in written and spoken languages (Biber et al., 2004; Cortes, 2004; Hyland, 2008a, 2008b) across various specific registers (Gardner & Davies, 2007; Liu, 2003, 2008), in various disciplines (Cortes, 2004; Durrant, 2017; Hyland, 2008a), in experts and students' writing (Chen & Baker, 2010), through various levels of proficiency in students' writing (Cortes, 2004, 2015; Chen & Baker, 2010; Staples, Egbert, Biber, & McClaire, 2013), in terms of the relationship between LBs' position and their communicative purposes (Cortes, 2013), and variations of LBs in L1 and L2 (Adel & Erman, 2012; Chen & Baker, 2010; De Cock, 2004; Pan, Reppen & Biber, 2016).

[DOI: 10.22034/jcals.1.1.67]

Regardless of the aforementioned findings of the previous studies, further research is needed to identify the characteristics of LBs specific to registers, genres, and disciplines. Furthermore, the frequency, form, and function of LBs specific to the conclusion section of applied linguistics RAs and their inconsistencies across international and Iranian journals are still under-explored areas of research. In other words, no study to date has neither identified LBs of the conclusion section of RAs in applied linguistics nor compared international and Iranian journals to find out whether or not LBs are used differently in their RAs. Therefore, the present study seeks to fill the gap of knowledge about the frequency, form, and function of LBs presented in the conclusion section of RAs in international and Iranian journals of applied linguistics and then to compare the results. Thus, three research questions guided this investigation:

1. What are the most frequent LBs and their forms and functions in the conclusion section of RAs in international journals of applied linguistics?

2. What are the most frequent LBs and their forms and functions in the conclusion section of articles in Iranian journals of applied linguistics?

3. What are the differences between the conclusion section of international and Iranian journals in terms of using LBs?

2. Methodology

The Corpus

Two corpora were compiled for this study including the applied linguistics RAs published in international journals and applied linguistics RAs published in Iranian journals. Given the impact factor scores and the publisher's popularity, only journals with the highest impact factor scores and also the most well recognized, globally acknowledged publishers have been selected to generate the former corpus. Consequently, the top five international applied linguistics journals were identified and selected for analysis. The selected journals include Applied Linguistics Journal (3.22), TESOL Quarterly (2.25), Journal of Language Learning (1.65), Modern Language Journal (2.78), and Language Teaching Research (2.08). The corpus is composed of 125 RAs published in the years 2015 to 2018. A number of twenty articles were randomly selected from each journal.

Due to the lack of a clear classification of Iranian journals, the expert opinion and availability of journals were considered criteria for developing the second corpus. Therefore, one hundred RAs were selected from five prestigious Iranian journals including Iranian Journal of Language Teaching Research (Urmia University), Issues in Language Teaching Journals (Allameh Tabataba'i University), Iranian Journal of Applied Linguistics (Kharazmi University), Applied Research on English Language (University of Isfahan), and Journal of Teaching Language Skills (Shiraz University). Twenty articles were randomly selected from each journal dated from 2015 to 2017. Table 1 presents a detailed description of two corpora in this study. The full list of articles is included in the appendix section.

[DOI: 10.22034/jcals.1.1.67]

Table 1. *Detail description of two corpora*

Corpus	International journals	Iranian journals
Journals	1. Applied Linguistics Journal 2. TESOL Quarterly 3. Journal of Language Learning 4. Modern Language Journal 5. Language Teaching Research	1. Iranian Journal of Language Teaching Research 2. Issues in Language Teaching Journals 3. Iranian Journal of Applied Linguistics 4. Applied Research on English Language 5. Journal of Teaching Language Skills
Year	2015-18	2015-17
Text type	RAs	Ras
No. of texts	125	100
Total number of words	52348	51856

In selecting the target research articles, hybrid headings (e.g., discussion and conclusion) were excluded because these sections provide different communicative functions (Lin & Evans, 2012).

Procedure & Data Analysis

The research articles were downloaded from the online version of the selected journals. Subsequently, the conclusion section of each article was extracted and converted into the plain text format for analysis. To perform accurate data processing, only the main text of the article was considered and the other parts, namely, header and footer for example, were removed from the files. Finally, the text was entered into the Antconc software, version 3.5.7, developed by Anthony (2006), for performing the analysis.

The analysis of this study is divided into three parts. The first concerns frequency, the second focuses on structural type, and the third one is related to functions of LBs.

Biber et al. (1999) set the frequency cut-off at 10 times per million words. In other words, multi-word sequences should at least be repeated 10 times in a corpus containing one million words to be regarded as LBs. This number ranges, however, from 20 to 40 occurrences per million words (Biber et al., 2004; Hyland, 2008b).

One of the primary issues in corpus linguistics is comparing the corpora of different sizes. This problem is dealt with through the process of normalization and the use of a concordance. Biber and Barbieri (2007) stated that a formula should be provided to normalize the count of LBs within a corpus. This process should be applied to corpora without an equivalent number of running words.

[DOI: 10.22034/jcals.1.1.67]

Chen and Baker (2010), who dealt with the implementation of the normalization process for an accurate comparison of LBs, explained that considering frequency and range criteria to establish the "optimum number of frequencies" is necessary (p. 32). These authors argued that determining the standard threshold for all corpora is not possible. They analyzed LBs in a 40,000-word corpus as opposed to an 80,000-word corpus. Considering the frequency threshold of 40 times per million words, the raw frequency is 1.6 for the former and 3.2 for the latter one (The raw frequency is calculated as follow: $\frac{40\times40000}{1000000} = 1.6$ and $\frac{40\times80000}{1000000} = 3.2$). To maintain an operational cut-off frequency, decimal values need to be rounded up or down.

Although the normalization process is still used as a method of identifying frequency in a corpus of less than one million words (Chen & Baker, 2016; Staples et al., 2013), it requires some further research (Chen & Baker, 2016). In this regard, Cortes (2002b) showed that more LBs are found in small corpora compared to the large ones because of the low-frequency requirement. Since there has not yet been an alternative method to compare smaller corpora, the use of the normalization procedure is regarded to be unavoidable in the present study. As a result, the frequency cut-off (according to normalization procedure formula $\frac{40\times50000}{1000000} = 2$) for both of the corpora used in the current study is equal to two.

Considering these criteria and following guidelines developed by Hyland (2008a), the bundles were classified according to the structural type. The bundles are then analyzed by functions proposed by Biber et al. (2004) as refined in Hyland (2008a). Finally, the results obtained from LBs in international and Iranian journals were compared, and a chi-square test was performed to assess whether the differences between the two corpora are significant or not.

3. Results

Overall Frequency of Lexical Bundles in Two Corpora

As mentioned earlier, the first purpose of this study was to analyze the frequency of LBs in international (INJs) and Iranian journals (IRJs). There were 109 and 192 four-word LBs in INJs and IRJs, respectively. The total numbers of bundles used in each corpus are represented in Figure 1. It is demonstrated that IRJs include more types and tokens (overall frequency of LBs used in a corpus) of bundles than INJs. In other words, Iranian authors utilized more bundles and used them more frequently than international writers.

[DOI: 10.22034/jcals.1.1.67]

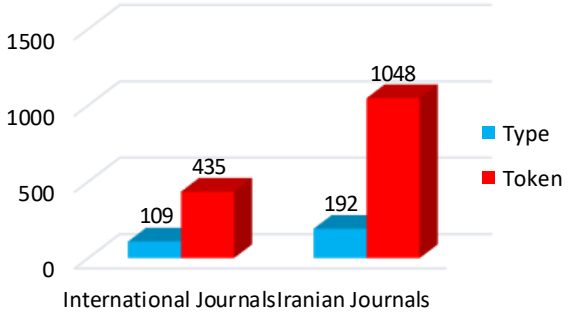

Figure 1. *Counts of types & tokens of bundles in corpora*

Structures of Identified Bundles

The structural framework of Hyland (2008a) revealed that his categories encompassed all structural correlations for all bundles. Table 2 indicates data concerning the structures of the identified bundles of both corpora.

Table 2. *Structural classification of lexical bundles in international &Iranian journals*

Structure	International Journals	Iranian Journals
Noun phrase + of	20 (18.35%)	34 (17.70%)
Other noun phrases	31 (28.44%)	50 (26.03%)
Prepositional phrase + of	18 (16.52%)	20 (10.42%)
Other prepositional phrases	19 (17.45%)	29 (15.10%)
Passive + prepositional phrase fragment	0 (0%)	16 (8.33%)
Anticipatory it + verb/adjective	8 (7.33%)	9 (4.70%)
Be + noun/adjectival phrase	3 (2.75%)	4 (2.10%)
Others	10 (9.16%)	30 (15.62%)
Total	109 (100%)	192 (100%)

In order to evaluate the structural features of LBs in this study, a chi-square test was conducted. The test results demonstrated that (see. Table 3) the differences between the two groups were significant at 0.05 ($p<0.05$).

Table3. *Chi-square test of structural analysis of lexical bundles in the two corpora*

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-square Test	14.708[a]	7	0.040
Likelihood Ration	20.036	7	0.005
Linear-by-Linear Association	2.077	1	0.150
N of Valid Cases	301		

[DOI: 10.22034/jcals.1.1.67]

Functions of Lexical Bundles in Two Corpora

The functional characteristics of LBs were analyzed using Hyland's (2008a) taxonomy. He categorized LBs in three major categories, focused on the research-, text-, and participant-oriented bundles and their corresponding sub-categories. Table 4 shows the overall function of LBs in two corpora.

Table 4: *Functions of lexical bundles in international &Iranian journals*

Function Type	Sub-function Type	International Journals	Iranian Journals
Research-oriented	Location	10 (9.43%)	4 (2.24%)
	Procedure	17 (16.03%)	35 (19.66%)
	Quantification	3 (2.83%)	9 (5.05%)
	Description	5 (4.71%)	8 (4.49%)
	Topic	5 (4.71%)	15 (8.42%)
Text-oriented	Transition signals	5 (4.71%)	7 (3.93%)
	Resultative signals	21 (19.85%)	41 (23.09%)
	Structuring signals	4 (3.77%)	7 (3.93%)
	Framing Signals	15 (14.15%)	22 (12.35%)
Participant-oriented	Stance features	9 (8.49%)	11 (6.17%)
	Engagement features	12 (11.32%)	19 (10.67%)
Total		106 (100%)	178 (100%)

A chi-square test was performed to identify significant differences between the functions of LBs across INJs and IRJs. According to Table 5, the result was not significant at the 0.05 level ($p<0.05$).

Table 5. *Chi-square test of lexical bundles in the two corpora*

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-square Test	10.632[a]	10	0.387
Likelihood Ration	10.493	10	0.398
Linear-by-Linear Association	0.005	1	0.943
N of Valid Cases	284		

4. Discussion & Conclusions

Frequency of Lexical Bundles

The analysis of this study showed that international authors used bundles less frequently than Iranian authors did. Like Esfandiari and Barbary's finding (2017), one reason might be attributed to the lack of structural unfamiliarity of LBs by Iranian researchers, which lacks any equivalent in the L2 language. For example, the LBs which start with *there* were found in IRJs (*there was no significant, there seems to be, there was a significant*) which were not used in INJs.

The second reason for using LBs more frequently by Iranian writers may be due to the L1 transfer of the author, which is in line with the result of Esfandiari and Barbary (2017). Iranian writers may have used a structure that seems untypical in English while trying to prepare their article. Paquot

[DOI: 10.22034/jcals.1.1.67]

(2013) referred to the L1 transfer impact, whereby the first language might lead learners to use LBs with untypical English patterns.

A number of 10 LBs containing '*this*' were found in IRJs (*this study was to, finding of this research, purpose of this study, this study was an, implication of this study, limitation of this study, the purpose of this, this study can be, this study could be, this study revealed that*), which were not found in INJs. In this respect, Adel and Erman (2012) suggested that the 'this' patterns are not very frequent in scholarly writing. In the conclusion section of the RAs, Iranian researchers applied "this" pattern followed by a head noun 'study' to demonstrate the resulting and structuring function**.** The International writers, however, used '*this*' pattern to explain the results as well as present the limitation and suggestions.

Culture might be the last reason for overusing LBs by Iranian writers. The previous study has revealed that formulaic sequences are usually culturally charged (Ma, 2009, p.127). More particularly, 165 four-word LBs were present in IRJs that did not even appear at lower frequencies in INJs. The Iranian authors used more LBs in their articles to express a meaning; "the longer, the more effective" is common practice in Persian (Esfandiari & Barbary, 2017, p. 27). Moreover, they seek to convey their opinions to readers as clearly as possible. The Iranian writers, for instance, used the longer form of the same bundle to make their sentences more effective (see extracts 1 and 2). Therefore, we can say that this result is consistent with those of Esfandiari and Barbary (2017).

- Indeed, ***this study suggests that*** such explicit attention to preservice teacher beliefs may be an excellent use of class time in teacher education (INJs).
- ***The outcomes of the present study suggested that*** manipulating task complexity along reasoning demand affected EFL learners writing accuracy provided that tasks were sequenced on the basis of cognitive complexity. (IRJs)

Structure of Lexical Bundles

To analyze the structural characteristics of LBs, Hyland's (2008a) taxonomy was employed which focused on eight structural types. As can be observed in Table 2, the primary distinction can be based on clausal and phrasal bundles. Phrasal bundles were the most recurring bundles in both corpora (mostly *other noun phrases* and *other prepositional phrases*). Some scholars, such as Biber et al. (2004), Biber (2006), Biber and Barbieri (2007), Chen and Baker (2010), Cortes (2004), and Hyland (2008a) found that the in academic writing, phrasal bundles are much more prevalent compared to the clausal one, which is consistent with the present study. Moreover, Pan et al. (2016) report that phrasal bundles are described mainly with a high concentration on information in academic writing. Hyland (2008a) also advocated that phrasal bundles play a crucial role in transmitting the knowledge of the

[DOI: 10.22034/jcals.1.1.67]

writer's arguments and focusing on significant qualities of RAs, resulting in phrasal bundles being applied in academic prose more than clausal bundles. This is also true in the current study, in which more than 80% of all bundles in INJs and 69% of bundles in IRJs were phrasal. Each one of these sections is explained in detail in the following paragraphs.

Noun Structure

Other noun phrases category, representing more than a quarter of all corpus bundles, is the most abundant structure among INJs (see Table 2). The IRJs are occupied, similar to the INJs, by noun structures that make up about half of all LB and, especially, by other substantive expressions. This structure consists of 50 bundles in IRJs, which is higher than those in INJs. This is in line with Biber et al. (1999), Byrd & Coxhead (2010), and Hyland's (2008a) findings and also highlights the notion that academic writing is "noun-centric" (Swales, 2008, p. v).

A careful examination of the LBs' concordances in this category indicated that the bundles were mostly used to convey meanings and generally match with *the … of* and *a… of* (i.e., the/a + noun [= NP] + of) framework that had been established earlier by Biber et al. (1999). The results were in agreement with Marco's (2000) subsequent conclusion that LBs in that category were implemented primarily to (1) indicate how a process or event occurred, (2) show the connection between the research items, and (3) evaluate research goals, materials, or participants if necessary.

In order to express their confidence in their interpretations of research results, Iranian authors more regularly used *the + Noun + of the* frame. These writers' high tendency for using these LBs reflects their lack of academic standards knowledge (Qin, 2014, p. 229). The concordance lines of the IRJs for the verbs that followed *the results of the* and *the findings of the* bundle had been thoroughly studied in order to better understand why they selected *the + Noun + of the*. The writers of IRJs used more direct verbs, including *reveal* (8 times), *suggest* (5 times), *indicate* (4 times), and *show* (2 times). International writers, however, used a far broader range of verbs, including *point to, provide, contribute, considered, suggest*, and *show*, to ease their assertions.

Considering the sub-category of *other noun phrases* in Table 2, a higher ratio of *other noun phrases* was determined in IRJs (50 bundles) compared to INJs (31 bundles). A reason why these bundles are high in frequency is that IRJ authors are recommended to select languages that are more native-like to their RAs. Due to the insufficiency of input (Schmitt, 2004) however, these scholars may use more or fewer LBs in their academic writing.

Prepositional Structures

This structure represented 34% and 26% of bundles in international and Iranian corpus, respectively. The preposition *in* was used in the majority of bundles in INJs that followed the prepositional structures (e.g., *in the field of, in the context of*). They were all primarily used to establish limitations for the presented arguments. On the other hand, bundles starting with the preposition *of* (e.g., *of the*

[DOI: 10.22034/jcals.1.1.67]

present study, of foreign language learning) were the most frequent bundles in IRJs and were used to express the ownership of different elements in this study.

The structure *in the + Noun + of* was also identified by Chen and Baker (2010) to be the major part of the *prepositional phrase + of* structures. In comparison to this, neither the international writers nor the Iranian writers seem to have recognized the importance of these prepositional expressions in their academic writing. Although the number of LBs following this category is similar in both corpora, the IRJs contain more tokens, which depends on *in the context of* bundle that was present 9 times/50,0000 words.

The numerous and diverse use of the prepositional structure in both companies demonstrates that the meaning of English preposition extends beyond what is known in the classroom context as an adverbial concrete (Byrd & Coxhead, 2010).

Verb Structures

Verb structures represent 9.19% and 12.11% of all LBs tokens and 10.09% and 15.10% of all types in the international and Iranian journals, respectively. Verbal constructs have greater structural diversity than noun structures, although they only have 80 keywords.

The most important point in the verbal analysis of LBs is about passive structure. Surprisingly, INJs, in contrast to IRJs, which include 16 passive bundles, contain no passive structure, and the result was statistically significant (see Table 3). This result is not consistent with recent findings (Biber et al., 1999; Chen & Baker, 2010; Hyland, 2008a). Most of these bundles, according to Biber et al. (1999), contain a passive voice verb followed by a prepositional phrase. The following section provides an example of LBs with verb phrase (including passive verb) from IRJs.

3) Further research ***is required to investigate*** the effects of task repetition on fluency in the production of speech acts. (IRJs)

Our study showed that the anticipatory *it* structure (such as *it is suggested that, it is possible that*) was used approximately equally in both corpora. As Hyland (2008a) states, anticipatory *it* is a way to mask authors' interpretations in advance. Anticipatory *it* pattern includes two different kinds, according to Biber (1999): the first one is followed by an adjective phrase (e.g., *it is important to*), and the second is followed by verbal sentences (usually passive voice) (e.g., i*t was found that*). International writers used mostly the first type of structure. There were seven bundles which followed the first type (*it is important to, it is also possible, it is also noteworthy, it is difficult to, it is imperative that, it is necessary to, it would be interesting*), and only one bundle followed the verb phrase (*it is hoped that*). In IRJs however, eight bundles followed passive verbs (*it can be concluded, it is suggested that, it should be noted, it was found that, it is hoped that, it should also be, it was revealed that, it is recommended that*).

[DOI: 10.22034/jcals.1.1.67]

Function of Lexical Bundles

As previously stated, the functional classification was done following Hyland's (2008a) framework. In this study, Table 4 indicates text-oriented bundles are the most frequent of the three primary functional categories in both corpora. This category includes 43% and 44% of LBs in INJs and IRJs, respectively. Research-oriented bundles, with 38% in INJs and 48% in IRJs, are the second frequent LBs in the two corpora. Both international and Iranian researchers used participant-oriented bundles less frequently than the other two; only 19% and 16% of bundles correspond to participant-oriented bundles in international and Iranian journals, respectively. These findings contrast to those reported by Hyland (2008a), who found that research-oriented bundles were the most common functional group in his science and technology corpora. He claimed that in the scientific ideology, the empirical is more frequent than interpretive as the emphasis is not on individual interpretations but actual demonstration and experimental findings. Research-oriented bundles thus provide scientific texts with a more realistic, laboratory-driven feeling. This contradiction can be explained by applying different disciplines in this study. Applied linguistics mainly includes the interpretation of findings and consequently, the number of research-oriented bundles will be reduced. Nonetheless, research-oriented bundles remained with reasonable efficiency to be recognized as an essential feature of academic writing in this study. In the next sections, each function will be thoroughly examined and probable causes are addressed.

Research-oriented Bundles

As Table 4 shows, research-oriented bundles provided a vast number of functions, ranging from representing time and location to the description of the issue. Procedural bundles are the most commonly used research-oriented bundles in both INJs and IRJs. Considering these kinds of bundles, one noteworthy result is that the tokens of this function are higher in IRJs than in INJs (139 versus 52 tokens). The topic-specificity of this special bundle function in particular and research-oriented bundles in general might be one reason for this. LBs such as *study examined the effects, through the use of, the quality of the,* and *and the validity of* may refer, due to subjects and goals variations between the two sets of researchers, to a particular method that was used by international authors, but not by Iranian writers. The reverse is true for the usage of extra procedural bundles in IRJs, such as *shed light on the, be aware of the,* and *the generalizability of the* due to the specificity of the topic.

The fairly less frequent research-oriented bundles – location, quantification, description, and topic - are used to characterize research subjects and contexts. According to Table 4, the international writers overused location bundles in their academic writing, and the result was statistically significant. The last type of research-oriented bundles, topic bundles, were related to the field-specific bundles and consisted of an expression that was used to express field-related concepts.

[DOI: 10.22034/jcals.1.1.67]

Text-oriented Bundles

For organizing the text, text-oriented bundles are employed which include four subtypes. Most LBs in both international and Iranian journals are associated with these functions. Hyland (2008a) finds text-oriented bundles as particularly indicative of soft knowledge-related areas that are more interpretive and less experimental, such as the applied linguistics. The same idea was obtained from the result of this study.

As Table 4 shows, the authors of IRJs implemented resultative signals (23.09%) more often in their writing than those of INJs (19.85%). The popularity of resultative signals may show these authors' lack of expertise and awareness of how to analyze research procedures and findings with these bundles.

A closer examination of IRJs revealed that the writers of IRJs used a few existential "*there*" patterns as a resultative function, while no "*there*" pattern was found in INJs. Biber et al. (1999) have indicated the statistical use of "*there*" bundles in RAs. However, for IRJs' writers, these bundles were utilized as a resultative function, since they are likely to misunderstand the function of these bundles.

This predominance of resultative signals is in accordance with the findings of Hyland (2008a), who identified significant use of these bundles in his biology corpus. In general, he emphasized that resultative signals play an important part in the rhetorical presentation of research and hence, are so frequent in hard science texts.

Framing signals, the second most common function in the text-oriented category, are another large number of bundles. Both international and Iranian journals had framing bundles similar to those found in Hyland (2008a); that is, they specify the limitations of arguments and center readers on a certain instance.

In looking at the differences between international and Iranian journals within framing bundles area, not much can be said about the frequency of these bundles (62 tokens in INJs versus 101 tokens in IRJs), since the international corpus contained few types of bundles overall (15 bundles in international and 22 bundles in Iranian journals). This can be attributed to variations in bundle use frequency. Iranian writers, for example, used *in the context of* nine times out of 50,000 words, while international scholars used it seven times in 50,000 words.

Two other text-oriented categories (transition signals and structuring signals) appear in smaller quantities than resultative and framing bundles. Most transition signal bundles had rather clear meanings (i*n addition it is, in other words the*). Most of the time, they were used to add or support a new argument or to contrast two situations.

The minimum number of text-oriented bundles was related to the structural group; only five bundles in INJs and seven in IRJs were included. The short length of the conclusion part examined in

[DOI: 10.22034/jcals.1.1.67]

this study might be responsible for the low number of bundles in this category. In his Ph.D. dissertation collection, Hyland (2008a) reported a significantly higher percentage of structuring bundles than in his corpus of research papers. He indicated that the longer the text is, the more time is needed to lead the readers through the text and show how a new argument or event ties to other previously discussed events or arguments.

Hyland (2008a) believes that structural bundles represent a high degree of dependency on graphical and tabular data (e.g., *are shown in table, as shown in figure*) and the requirement to refer to such data in the presentation of events or arguments. However, no such bundles were found in the corpora of the current study.

Participant-oriented Bundles

The last major category of functions refers to participant-oriented bundles. In both international and Iranian journals, participant-oriented bundles are less frequent than the other two groups. Furthermore, in both international and Iranian journals, engagement bundles were more common than stance bundles. This is in line with Hyland's (2008a) results, which showed that engagement bundles dominated his science corpora. He argued that the considerable prevalence of these bundles demonstrated how important it is for hard science fields to ensure that methods and outcomes are understood correctly.

The first subcategory, stance bundles, in the current study, was utilized for transferring meaning and function like those found in Cortes' biology corpus (2004). She claimed that the majority of the bundles in this group suggest a certain level of probability and also convey "a degree of tentativeness" to the arguments or findings being provided (p. 410). Regarding INJs and IRJs, nine (8.49%) and 11 (6.17%) types of LBs were identified as stance bundles, respectively. Of the nine types identified in INJs, seven appeared three times, and only two have more than three occurrences: *has the potential to* and *it is also possible* (4 occurrences). The same is true for IRJs. Most of the bundles appeared three or four times in IRJs, and only one bundle occurred five times: *seems to be a*.

Hyland (2012) indicated that due to "cultural preferences for noninterventionist stance bundles", Chinese students incorporated less stance in their doctoral dissertation. This result, however, is not consistent with the result of the present study because more stance bundles are used in IRJs.

The other participant-oriented sub-category is engagement features, which is a means by which readers may acknowledge and comprehend recognized disciplinary knowledge (Hyland, 2005). This was more prevalent in the IRJs, comprising 63% of participant bundles, as compared to just 57% in the INJs.

A detailed investigation of participant-oriented bundles showed that first, this functional category contains more hedges than the other two categories, and second, the writers of IRJs adopted

[DOI: 10.22034/jcals.1.1.67]

more hedging-based LBs (e.g., *it is possible that, can be considered as, there seems to be*) compared to writers of INJs. The INJs consisted of only three hedging-based LBs, including *it is also possible, future studies should consider,* and *we believe that the*. The authors of IRJs have therefore not only chosen a higher proportion of LBs with hedges, but also a wide variety of lexical phrases (e.g., *can, possible, seem*). Hyland (2008a) also agrees with this result, emphasizing the formulaic structure of engagement features and how they contribute to the accuracy that marks scientific writing.

The purpose of this study is to contribute to the field of applied linguistics by tackling issues like what the differences between international and Iranian journals are in terms of frequency, form, and function of LBs found in the conclusion section of applied linguistics RAs.

In terms of the first criterion, which focused on the total bundle frequency, Iranian writers utilized a broader range of bundles with a higher frequency than international writers. When bundles are shared between international and Iranian journals and compared to the other three studies, the four groups use them very differently.

Structurally, noun phrase bundles formed the largest category both in international and Iranian journals. The major structural differences were found in the two corpora. The writers of INJs did not use any passive structures, while Iranian writers significantly utilized passive structures in their academic writings. This difference made the structural differences between international and Iranian journals significant.

International and Iranian writers also differed on the functional distribution of LBs. The writers of both international and Iranian journals used text-oriented bundles more frequently than the other two categories. However, they used location, procedure, and topic in the research-oriented section as well as the resultative signals in the text-oriented section differently. The only sub-category which was significantly different between the two corpora was location. However, the overall functional analysis of LBs in the two corpora was not significant.

LBs can differentiate highly professional journals from moderate ones in some respects. For example, it was found that passive structure is overused by Iranian writers, which are considered a deviation from the international writers' conventions. It can be concluded that LBs cannot be entirely considered a criterion for distinguishing the degree of professionalism in academic writing.

The outcome of this research, from a pedagogical perspective, may contribute to the development of more beneficial materials for academic writing. Such results may also be of significance to the experts in the field of applied linguistics and, more notably, beginner writers in the discourse community, eager to become professional members of their specific community.

This paper has extended the current understanding of lexical bundles. However, additional investigations on the use of lexical bundles, particularly in Iranian contexts, are necessary in the

[DOI: 10.22034/jcals.1.1.67]

future. To compare each of the bundles identified in this study with those previously discussed in the literature; it would be interesting to check if the Iranian writers employed the same LBs in the same way, and for the same purposes and function. Moreover, the present study mainly focused on LBs in texts composed by professional authors and, hence, tells only one-half of the story. The kinds of bundles employed by beginner authors as well as comparing them to those used by expert writers in the field of applied linguistics are the other half of the story. Thus, the next logical step might be to investigate the use of LBs in the learners' writings and to compare them with the ones discovered in this study. Additionally, the present study focused exclusively on four-word bundles. Expanding the length of LBs for analysis of international and Iranian journals could reveal different keywords and LBs and consequently, a different result, which can be the purpose of further research. Finally, the researcher selected the conclusion section of applied linguistics RAs for analysis. Further researches can investigate how LBs are used in different sections of applied linguistics academic writings, such as introduction, methodology, etc.

Authors' Contributions

All authors contributed significantly to the research process.

Declaration

We declare that this manuscript is original and has not been submitted to any other journal for publication

Transparency Statements

The authors affirm that the data supporting the findings of this study are available within the article. Any additional data can be obtained from the corresponding author upon reasonable request.

Acknowledgements

We would like to express our gratitude to all individuals helped us to do the project.

Declaration of Interest

The authors report no conflict of interest.

Funding

The authors do not have any financial or non-financial competing interests.

Ethical Consideration

This manuscript adheres to the ethical guidelines provided by the Committee on Publication Ethics (COPE) for ensuring integrity and transparency in the research publication process.

[DOI: 10.22034/jcals.1.1.67]

References

Ädel, A., & Erman, B. (2012). Recurrent word combinations in academic writing by native and non-native speakers of English: A lexical bundles approach. *English for specific purposes*, *31*(2), 81-92.

Anthony, L. (2006). Developing a freeware, multiplatform corpus analysis toolkit for the technical writing classroom. *IEEE Transactions on Professional Communication*, *49*(3), 275-286.

Biber, D. (2006). *University language: A corpus-based study of spoken and written registers* (Vol. 23). John Benjamins Publishing.

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

Biber, D., & Conrad, S. (1999). Lexical bundles in conversation and academic prose. *Language and Computers*, *26*, 181-190.

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). Longman grammar of spoken and written language. *Harlow: Pearson Education*.

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*, *5*(5), 31-64.

Chen, Y. H., & Baker, P. (2010). Lexical Bundles in L1 and L2 Academic Writing. *Language Learning and Technology*, *14*(2), 30-49.

Chen, Y. H., & Baker, P. (2016). Investigating criterial discourse features across second language development: Lexical bundles in rated learner essays, CEFR B1, B2 and C1. *Applied Linguistics*, *37*(6), 849-880.

Cortes, V. (2002b). Lexical bundles in academic writing in history and biology. (Unpublished doctoral dissertation). Northern Arizona University, United States.

Cortes, V. (2004). Lexical bundles in published and student disciplinary writing: Examples from history and biology. *English for specific purposes*, *23*(4), 397-423.

Cortes, V. (2006). Teaching lexical bundles in the disciplines: An example from a writing intensive history class. *Linguistics and Education*, *17*(4), 391-406.

Cortes, V. (2013). The purpose of this study is to: Connecting lexical bundles and moves in research article introductions. *Journal of English for academic purposes*, *12*(1), 33-43.

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.

Cortes, V., Jones, J., & Stoller, F. (2002, April). Lexical bundles in ESP reading and writing. In *TESOL conference, Salt Lake City, Utah, USA*.

Coxhead, A. (2008). Phraseology and English for academic purposes. *Phraseology in language learning and teaching*, 149-161.

De Cock, S. (2003). *Recurrent sequences of words in native speaker and advanced learner spoken and written English: A corpus-driven approach* (Doctoral dissertation, UCL-Université Catholique de Louvain).

[DOI: 10.22034/jcals.1.1.67]

De Cock, S. (2004). Preferred sequences of words in NS and NNS speech. *Belgian Journal of English Language and Literatures (BELL), New Series, 2*(2), 225-246.

Durrant, P. (2017). Lexical bundles and disciplinary variation in university students' writing: Mapping the territories. *Applied Linguistics, 38*(2), 165-193.

Esfandiari, R., & Barbary, F. (2017). A contrastive corpus-driven study of lexical bundles between English writers and Persian writers in psychology research articles. *Journal of English for Academic Purposes*, 29, 21–42. https://doi.org/10.1016/j.jeap.2017.09.002

Gardner, D. (2007). Validating the construct of word in applied corpus-based vocabulary research: A critical survey. *Applied Linguistics, 28*(2), 241-265.

Gardner, D. E. E., & Davies, M. (2007). Pointing out frequent phrasal verbs: A corpus-based analysis. *TESOL QUARTERLY, 41*(2), 339–359.

Granger, S. (1998). Prefabricated patterns in advanced EFL writing: Collocations and formulae. *Phraseology: Theory, analysis, and applications*, *145*, 160.

Granger, S. (2014). A lexical bundle approach to comparing languages: Stems in English and French. *Languages in Contrast*, *14*(1), 58-72.

Granger, S., Hung, J., & Petch-Tyson, S. (Eds.). (2002). *Computer learner corpora, second language acquisition, and foreign language teaching* (Vol. 6). John Benjamins Publishing.

Howarth, P. (1998). Phraseology and second language proficiency. *Applied linguistics*, *19*(1), 24-44.

Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge University Press.

Hyland, K. (2005). *Metadiscourse: Exploring interaction in writing*. A&C Black.

Hyland, K. (2008a). As can be seen: Lexical bundles and disciplinary variation. *English for specific purposes*, *27*(1), 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). *Disciplinary identities: Individuality and community in academic discourse*. Cambridge University Press.

Hyland, K., & Hamp-Lyons, L. (2002). EAP: Issues and directions. *Journal of English for academic purposes*, *1*(1), 1-12.

Jablonkai, R. (2010). English in the context of European integration: A corpus-driven analysis of lexical bundles in English EU documents. *English for Specific Purposes*, *29*(4), 253-267.

Jaworska, S., Krummes, C., & Ensslin, A. (2015). Formulaic sequences in native and non-native argumentative writing in German. *International Journal of Corpus Linguistics*, *20*(4), 500-525.

Lin, L., & Evans, S. (2012). Structural patterns in empirical research articles: A cross-disciplinary study. *English for Specific Purposes, 31*(3), 150-160.

Liu, D. (2003). The most frequently used spoken American English idioms: A corpus analysis and its implications. *Tesol Quarterly*, *37*(4), 671-700.

Liu, D. (2008). Linking adverbials: An across-register corpus study and its implications. *International Journal of Corpus Linguistics*, *13*(4), 491-518.

[DOI: 10.22034/jcals.1.1.67]

Ma, G. H. (2009). Lexical bundles in L2 timed writing of English majors. *Foreign Language Teaching and Research*, *1*, 54-59.

Marco, M. J. L. (2000). Collocational frameworks in medical research papers: A genre-based study. *English for specific purposes*, 19(1), 63-86.

Nesselhauf, N. (2005). *Collocations in a learner corpus* (Vol. 14). John Benjamins Publishing.

O'keeffe, A., McCarthy, M., & Carter, R. (2007). *From corpus to classroom: Language use and language teaching*. Cambridge University Press.

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.

Paquot, M. (2013). Lexical bundles and L1 transfer effects. *International Journal of Corpus Linguistics*, *18*(3), 391-417.

Qin, J. (2014). Use of formulaic bundles by non-native English graduate writers and published authors in applied linguistics. *System*, *42*, 220-231.

Schmitt, N. (Ed.). (2004). *Formulaic sequences: Acquisition, processing, and use* (Vol. 9). John Benjamins Publishing.

Sinclair, J. M. (Ed.). (2004). *How to use corpora in language teaching* (Vol. 12). John Benjamins Publishing.

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.

Swales, J. M. (2008). Foreword. In D. Belcher & A. Hirvela (Eds.), *The oral-literate connection: Perspectives on L2 speaking, writing, and other media interactions* (p. v-viii). Ann Arbor: University of Michigan Press.

Wray, A. (2000). Formulaic sequences in second language teaching: Principle and practice. *Applied linguistics*, *21*(4), 463-489.

Wray, A. (2005). *Formulaic language and the lexicon*. Cambridge University Press.

[DOI: 10.22034/jcals.1.1.67]