РАЗДЕЛ II. ПРИКЛАДНАЯ ЛИНГВИСТИКА
SECTION II. APPLIED LINGUISTICS

UDC 81'33 **DOI: 10.18413/2313-8912-2024-10-1-0-4**

Atta A. Ahmed[1]
Aida Ariannejad[2]

A comparative study of lexical bundles in linguistics and biology Ph.D. dissertations

[1] Tishk International University-Sulaimani
Qularaisi, Sulaymaniyah, 46001, Kurdistan, Iraq
E-mail: atta.ahmed@tiu.edu.iq

[2] World Peace University
Şht. Kemal Ali Ömer St. No:22 Yenişehir, Nicosia, 99010, Turkish Republic of North Cyprus, Mersin 10, Turkey
E-mail: aida.ariannejad@wpu.edu.tr

Received 17 December 2022; accepted 12 February 2024; published 30 March 2024

Abstract. Lexical bundles are considered as one of the main rhetorical features in academic genres. The appropriate use of these features adds to the coherence and naturalness of the texts. Recently, genre analysis studies tend to investigate the similarities and differences across different academic disciplines and the effects of the first language of writers on the employment of these linguistic features. There are limited studies, however, in the field with a mere focus on the use of lexical bundles in PhD dissertations and the similarities and differences on the employment of these features among native and non-native English writers. In this regard, the current study, following a comparative corpus-based approach, investigated the use of lexical bundles in English PhD dissertations written by native English-speaking and non-native Kurdish-speaking writers across the two disciplines of biology and linguistics. All the compiled dissertations were selected from the ones published between 2010 to 2020 in British universities to keep the compiled corpora comparable. In the next phase, the distribution, linguistic structures, and functions of bundles used in introduction and literature review sections of the compiled dissertations were analyzed using WordSmith 6th edition concordancing software. The analysis provides a list of the most frequently used lexical bundles in each scientific field and among the two groups of writers. The structural analysis showed that noun phrases and prepositional phrases were the most frequently used bundles between both groups of writers and of disciplines. It was also found that passive structures were commonly used in biology dissertations. The functional analysis revealed that non-native Kurdish writers tended to use more research-oriented and text-oriented bundles in comparison with their native counterparts. It was suggested that the rhetorical differences and similarities between the two groups of writers and the disciplines could be attributed to the experience or proficiency levels of writers and the conventions, knowledge construction principles and research approaches of each

academic discourse community. Pedagogically speaking, it is discussed that teaching discursive features and including language teaching tasks drawn from authentic texts such as text analysis tasks, focused tasks, extended writing tasks enable EAP learners to learn how to employ these linguistic features effectively to improve the flow of information, coherence and fluency in their academic texts.

Keywords: EAP; Lexical bundles; PhD dissertations; Cross-disciplinary; Native and non-native academic writers; Corpus-based analysis

How to cite: Ahmed, A. A. and Ariannejad, A. (2024). A comparative study of lexical bundles in linguistics and biology Ph.D. dissertations, *Research Result. Theoretical and Applied Linguistics*, 10 (1), 47-60. DOI: 10.18413/2313-8912-2024-10-1-0-4

Introduction

Recently, a number of academic genre analysis studies focus on a new category of word sequences, i.e., lexical bundles, which are defined as fixed sets of words which frequently occur in natural language use (Baker & Ellece, 2011). The term "lexical bundles" was first coined by Biber, Johansson, Leech, Conrad, and Finegan (1999). The term refers to expressions such as *at the same time*, *the results show that*, *could be suggested that*, and *it is necessary to* which occur reasonably frequently in a specific register. For a frequent cluster of words to be considered as a distinctive lexical bundle of a specific genre, it must occur at least twenty times in a one-million-word corpus and must be used in at least five or more texts to control factors such as the idiosyncratic styles of writers and repetitions (Hyland, 2008a).

Hyland (2008a) states that lexical bundles are key rhetorical features used in academic contexts since their appropriate employment and familiarity with the structures and functions of these expressions enable writers to add to the naturalness of their text. The appropriate employment of these expressions could act as a key feature in distinguishing an experienced writer from a novice one in an academic discourse community. Thus, raising the awareness of the novice academic writers of the important role of lexical bundles in increasing their writing fluency and teaching them the way such clusters are used and occur in specific genres can assist them to develop their discourse competence and their knowledge regarding the rhetorical norms and preferences of their target academic discourse community. In a broader sense, having knowledge of lexical bundles would act as a facilitative and motivating element in the linguistic performance of second and foreign language learners (Jones & Haywood, 2004).

In the earlier studies on lexical bundles, Biber et al. (1999) explored the principal structures of bundles used in various academic discourse and Cortes (2004, 2008) highlighted the salient role lexical bundles play in improving the coherence and organization of the texts through serving a broad range of discursive functions. Scott and Tribble (2006) and Biber (2006), have also emphasized the reliance of various academic discourses on a diverse repertoire of lexical bundles. Thus, in the attempts to organize the abundant number clusters found in various academic discourse, researchers classified them into structural and functional categories (Biber, Conrad, & Cortes, 2004). Such classifications provided a more comprehensive view of how these features are used across various academic discourse.

There have been a number of investigations on the structures and functions of lexical bundles across different disciplines (Hyland, 2008a, 2008b; Yin & Li, 2021). Others compared the use of bundles between the two groups of native and non-native academic writers (Ädel & Erman, 2012; Ucal, 2017). Some studies investigated the employment of these expressions across various academic genres such as research articles (Candarli & Jones, 2019; Cao, 2021),

or textbooks (Gholaminejad, 2021). Yet, there are still specific contexts, disciplines and genres which need to be further investigated. For instance, there are limited number of investigations with a mere and in-depth focus on other important academic genres such as PhD dissertations. PhD dissertations are the most distinguished academic genre produced by graduate students (Jalili, 2013) and as Hyland (2008b) stated PhD dissertations "... carry the burden of assessment and determine future life chances, but with different expectations for particular forms of argument, cohesion, and reader engagement" (Hyland, 2008b: 50). There is no doubt that this area of research deserves to be explored further in future lexical bundles studies.

Considering the above-mentioned issues, this study adopts a comparative approach to investigate the employment of lexical bundles in PhD dissertations written in English by the two groups of Native English-Speaking (NES) and Non-Native Kurdish-Speaking (NNKS) writers in the fields of biology and linguistics. In order to contribute to the body of this paper, we referred to a study that is conducted by Ahmed (2021) on *The Use of Lexical Bundles in English PhD Dissertations Written by Native English-speaking and Non-native Kurdish-speaking Writers across the Two Disciplines of Biology and Linguistics*. In fact, the current paper is a compact, revised and enhanced report of the original study. In this paper, the fields of biology and linguistics were selected to represent the two areas of hard sciences and soft sciences, respectively. The comparison between the biology and linguistics texts in the current research would make it possible to generalize the results considering the epistemological and knowledge construction differences between the two contrasting streams of soft and hard sciences. Hyland (2009) distinguishes the two domains arguing that in hard sciences such as Chemistry or biology the knowledge is constructed on empirical basis and experimental methods for a concentrated group of readers whereas in the soft sciences, such as linguistics, or psychology, knowledge construction is based on explicit interpretation of the data and discursive arguments to support the claims for the relatively more homogeneous and wider groups of audience. A further unique aspect of this study is the attempt to explore the rhetorical preferences of non-native Kurdish-speaking English writers in the employment of lexical bundles. To the best knowledge of the researchers, there are limited number of academic genre analysis studies on the rhetorical patterns and norms used by this group of writers. Although the current study aims to investigate the rhetorical patterns used by Kurdish-speaking writers in English text, it is believed that such research sheds light on a number of linguistic features used by this group which are shaped under the influence of their first language, i.e. Kurdish. Taking the aforementioned issues into consideration, the present study seeks to explore the most frequent lexical bundles and their main structures, and functions in PhD dissertations written in English by the two groups of NES and NNKS writers in the fields of biology and linguistics.

Literature Review

There is a growing body of research on lexical bundles and their structural and functional categories on the literature. A number of these studies investigated the employment of lexical bundles in and across disciplines. In one of the most comprehensive studies in the field, Hyland (2008a) studied the frequency, structure and functions of lexical bundles used in various academic genres across electrical engineering and biology, as the representative fields of hard sciences and business studies and applied linguistics, as the representative fields of soft sciences. The results revealed that in comparison to other fields, academic writers of electrical engineering have a higher tendency to use prefabricated structures in knowledge construction and the reports of their findings. The structural analysis of the lexical bundles revealed that noun-phrases following an *of*-fragment were the most frequently used bundle structure across the

four disciplines. The results also showed the tendency of hard science writers of biology and electrical engineering fields to use more passive bundles followed by a prepositional phrase, whereas, in soft sciences, writers tend to use a wide variety of prepositional phrases. It was suggested that the differences in the use of lexical bundles among the disciplines were to a great extant attributed to differences in knowledge construction methods and persuasive strategies used by the discourse community members of each discipline. In his investigation on the functions of lexical bundles, following Halliday's (1999) metafunction classifications, Hyland (2008a) classified the functions of lexical bundles to three categories of research-oriented, participant-oriented, and text-oriented functions. Firstly, research-oriented bundles with ideational roles assist the researcher to represent their real-world activities, and experiences. Writers use research-oriented bundles to give information about the location of the research, the procedures followed in the study, and the issues related to quantification and description of the data or the context. Secondly, participant-oriented activities serve an interactional role and enable writers to express their own stance and attitudes towards the propositions and to directly involve the readers in the arguments. Finally, text-oriented bundles take textual roles to assist writers to organize their texts through establishing transitions between the arguments, marking causative relations, referring to information elsewhere in the text, and introducing limiting conditions for the arguments. In line with Hyland's study, researchers such as Byrd and Coxhead (2010) investigated the diversity, frequency, structures and functions of bundles used across a number of disciplines. In a more recent study, Yin and Li (2021) compared the employment of bundles in research articles published in the fields of finance, accounting, biology, and applied linguistics. Their findings revealed that even contextually closer fields of finance and accounting differ significantly in the use of resultative clusters, although still similar in the topic-related ones. It was also argued that business articles contained more bundles in comparison with applied linguistics and biology articles, suggesting that the phraseological behavior of business papers might be stronger than other compared disciplines.

Research has also shown that genres differ in the employment of lexical bundles. A glance at the lexical bundle studies with a focus on academic genres reveals that such investigations mainly focus on the use of bundles in research articles (Candarli & Jones, 2019; Cao, 2021; Hernández, 2020; Jalali & Moini, 2018). In a recent cross-generic study, Gholaminejad (2021) investigated bundles used in applied linguistics textbooks and research articles. The results showed that research articles included more research-oriented and text-oriented bundles than the textbooks. Hyland (2008b) compared PhD dissertations, master theses, and research articles. The findings revealed that master theses included more bundles in contrast with the other genres. The study reported functional similarities between research articles and PhD dissertations in the higher frequency of text-oriented bundles which was discussed to possibly be the result of higher proficiency or experience of their writers in comparison to master theses produced by less proficient writers. More recently, Kashiha and Heng (2015) compared the structures of bundles used in university lectures across the disciplines of chemistry and politics and Zare and Valipour (2021) merely focused on the grammatical structures of lexical bundles used in research articles in the field of chemistry. In another research, Nasrabady, Shirvan and Golparvar (2020) focused on the bundles used in applied linguistics articles published in high stake journals.

Another group of studies in the field explored the possible similarities and differences on the use of lexical bundles among native and non-native writers. The findings revealed that there are differences among the two groups of writers in the

employment of these features (Amirian, Ketab & Eshaghi, 2013; Gil & Caro, 2019). Chen and Baker (2010) studied bundles used in English essays written by English-speaking native expert, English-speaking native non-expert and non-native Chinese-speaking novice writers structurally and functionally. The results revealed basic similarities in the use of lexical bundles among the two groups of novice writers, i.e., native English-speaking and non-native Chinese-speaking, specifically, in their reliance on verb-phrase based bundles which was discussed to be due to their lower proficiency level. The native English-speaking experienced writers, on the other hand, were found to use more noun-phrase bundles and referential expressions. In another study, Appel and Wood's (2016) explored the use of bundles between low-level and high-level academic essays of non-native writers, from the Canadian language assessment center. Their findings revealed that the low-level writers tend to use more sequencing bundles, as well as bundles with stance and discourse-organizing functions in comparison to the high-level writers who used relatively more referential bundles. More recently, Yakut et al. (2021) compared the use of lexical bundles in English PhD dissertations written by native English-speaking writers and non-native Turkish speaking writers. The data included dissertations published in English language related fields between 2010 to 2019. They found that non-native Turkish writers tend to use more lexical bundles in comparison with their native counterparts which might be associated with the tendency of postgraduate writers to gain credibility from their expert readers and editors. The structural analysis also revealed that the two groups were different in the employment of noun phrases and prepositional phrases, and that these two structures were the most frequently used ones across both native and non-native corpora. Moreover, text-oriented functions were found to be the most commonly used bundles among both groups of writers. The reason for such tendencies were discussed to be mainly related to the awareness of postgraduate writers to meet the expectations and norms of their academic discourse community and their willingness to show themselves as competent members of the community.

Methodology

The Corpus

The corpus compiled in the present study included biology and linguistics PhD dissertations written in English. The corpus included four sub-corpora: biology dissertations written by native English-speaking writers, linguistics dissertations written by native English-speaking writers, biology dissertations written by non-native Kurdish-speaking writers, and linguistics dissertations written by non-native Kurdish-speaking writers. Moreover, the compiled corpus only included the introduction and literature review sections of the dissertations. Analyzing the rhetorical features of the introduction sections of dissertations is conducive since it is in this section that writers discuss the significance and the novelty of their work. Writers in the literature review sections provide a deeper value of the current study by constructing the body of knowledge which is valuable and persuasive for the academic community members (Hyland, 2009).

As it is shown in Table 1, the biology sub-corpora, comprised of English dissertations written by the Native English-Speaking (NES) and Non-Native Kurdish-Speaking (NNKS) writers, in total, includes 177,391 words and the linguistic sub-corpora, comprised of dissertations written by the NES and NNKS writers includes 285,587 words, in total. Moreover, the NES sub-corpora comprised of linguistics and biology dissertations written by native English writers included 204,402 words. The NNKS sub-corpora of linguistics and biology dissertations written by non-native Kurdish writers included 318,576 words. The whole corpus includes 462,978 words.

Table 1. Word Counts of the Compiled Sub- Corpora

	Biology Sub-corpora	Linguistics Sub-corpora	Total word count
NES sub-corpora	71,074	133,328	204,402
NNKS sub-corpora	106,317	152,259	258,576
Total	177,391	285,587	462,978

Corpus compilation procedure

It is worth noting that all the dissertations were downloaded from the British Library e-theses online service (https://ethos.bl.uk) and therefore they were all published in the universities of the United Kingdom. All the dissertations were published between 2010 and 2020. In addition, the English-speaking writers had English names and surnames and were affiliated with the British universities. In the same way, the Kurdish-speaking writers had Kurdish names and surnames. The information in the acknowledgement sections and the contexts of the studies were taken into consideration to ensure that the writers were Kurdish. Finally, in the corpus compilation process, following Sinclair's (1991) clean-text policy, the tables, footnotes, endnotes, figures, and diagrams were removed from the dissertations.

Data Analysis

In order to find the most frequently employed bundles in the compiled sub-corpora, the concordancing software of WordSmith 6th edition[1] was used. This study mainly reports four-word bundles found by the concordance, since four-word bundles are more frequently used in the academic texts than the five-word bundles. Also, compared to three-word bundles, four-word bundles serve a broader variety of structures and functions (Cortes, 2008; Hyland, 2008a; Jalili, 2013). The cut-off points for the four-word bundles extracted from the corpus was set at the raw frequency of 10. This is because considering the sizes of the native/non-native and cross-disciplinary sub-corpora compiled in this study, whose average word count are around 250,000 words, the frequency cut-off point was selected to be 40 per a million word or 10 in each 250,000 words.

In the analysis stage, the concordance data were manually annotated, taking into account the co-texts where the chunks were used, to check the overlapping bundles and to study the structure and functions of the bundles. At this phase, the bundles containing proper nouns such as *Kurdistan region of Iraq*, bundles directly associated with specific topics in biology or linguistics such as *the RNA exit channel*, *language variation and change*, and discipline related bundles such as *stem of the verb,* and *in the nervous system* were excluded. The reason to eliminate such bundles is to minimize the effects of the specific dissertation topics on the employed bundles and to control the number of repetitive bundles used in the same texts (Ädel and Erman, 2012).

In the next phase, the frequency of the four-word bundles was compared across disciplines and among the two groups of native English and non-native Kurdish writers. The bundles were later classified structurally, following Biber's et al. (1999) structural taxonomy. The functional analysis of the bundles was conducted based on Hyland's (2008a, b) functional classification of research-oriented, text-oriented, and participant-oriented functions, which was discussed earlier in this paper.

Results and discussions

Frequency and distribution of lexical bundles across the sub-corpora

In order to investigate the similarities and differences in the use of lexical bundles by the two groups of Native English-Speaking (NES) and Non-Native Kurdish-Speaking

[1] Scott, M. (2015). WordSmith tools (Version 6) [Computer software], Oxford University Press, Oxford, UK.

(NNKS) writers in the fields of biology and linguistics, the compiled corpora were analyzed from three different perspectives, first, the distribution and frequency of the bundles, second their main linguistic structures, and finally the functions of these features across various contexts. Accordingly, Table 2 shows the types of bundles (the number of concurrent bundles) and tokens (the frequency of each bundle) found in the compiled corpus. In total 1902 tokens and 119 different bundle types were found across the whole corpus.

Table 2. Types and Tokens of Lexical Bundles

	Biology Sub-corpora		Linguistics Sub-corpora		Total	
	Type of bundles	Tokens	Type of bundles	Tokens	Type of bundles	Tokens
NES sub-corpora	6	83	36	569	42	652
NNKS sub-corpora	15	232	63	1018	78	1250
Total	21	315	99	1587	119	1902

The cross-disciplinary comparison between biology and linguistics dissertations revealed that biology dissertations included 315 tokens and 21 different bundle types which is fewer than the frequency of bundles in linguistics dissertations, including 1587 tokens and 99 cluster types. Moreover, it was found that NES writers used 42 types of bundles with the frequency of 652 while their NNKS writers used more varieties of bundles, 78 types, with higher frequency of 1250 tokens.

The cross-disciplinary result of this study is consistent with Cortes (2004) and Hyland (2008a) studies. They also found that the frequency of bundles in biology academic texts to be fewer than other disciplines. The reason could be associated with the reliance of biology writers on a wide range of field-specific, technical words and expressions and their tendency to use more naming and coding than other disciplines. Moreover, the higher frequency of bundles in NNKS sub-corpus could indicate the lower proficiency level of non-native Kurdish writers. In this regard, Hyland (2008b) argued that novice writers are more dependent on formulaic chunks and their genre is more phrasal in comparison to experienced or native writers.

Table 3 shows the 10 most frequent bundles used in each sub-corpus. As it is shown, *On the other hand* was repeated in three sub-corpora and was found to be the most frequently used bundles in the whole corpus, with the frequency of 96. Moreover, it was found that both native and non-native writers of the biology dissertations used various phrases with the stem noun *presence,* such as *the presence of the*, *by the presence of*, *in the presence of*. The recurrent bundles of biology sub-corpora suggest that the biology writers, in the introduction and literature review sections of their dissertations, prefer to provide explanations for the field-related processes and to discuss the existence or non-existence of various elements in such processes using varieties of bundles.

In addition, it was shown that the native and non-native writers of linguistics dissertations shared some common bundles such as *on the other hand, the use of the, on the basis of,* and *in the case of*. This might probably indicate the attempts of linguistics writers to persuade their readers through creating plausible connection between different aspects of their arguments. Moreover, the comparison showed a limited number of shared bundles, such as *as a result of,* in linguistics dissertations written by the two groups. It seems that the selection of lexical bundles by the writers in this field

depends more on the rhetorical patterns of their academic discourse and the contents they discuss, rather than the effects of their first language. It could also be due to the limitations of the study in compiling a rather small corpus.

Table 3. Most Frequent 10 4-Word Bundles in each Sub-Corpus

NES Biology sub-corpus	NES Linguistics sub-corpus	NNKS Biology sub-corpus	NNKS Linguistics sub-corpus
has been shown to	as well as the	**as a result of**	**on the other hand**
as a result of	the extent to which	an increase in the	**the use of the**
the presence of the	**the use of the**	a wide range of	**on the basis of**
the structure of the	**on the basis of**	in the presence of	**in the case of**
by the presence of	**on the other hand**	is one of the	might be argued that
the majority of the	it is possible to	**on the other hand**	with regard to the
in the regulation of	**in the case of**	**has been shown to**	**as a result of**
the results of the	at the same time	in the absence of	the representation of the
a wide range of	it should be noted	in the process of	the number of the
the quality of the	in terms of the	an important role in	the fact that the

The structural analysis of lexical bundles

In Tables 4 and 5 the findings are shown in NES, NNKS, biology and linguistics classifications to provide a clear view of the patterns used in each sub-corpus. The NES sub-corpus includes the biology and linguistics dissertations written by NES writers. The NNKS follows the similar discipline classification. Moreover, each of the biology and linguistic sub-corpora includes dissertations written by both NES and NNKS writers.

Table 4. Main Structures of Bundles in each Sub-Corpus (%)

Structure	NES	NNKS	Biology	Linguistics	Total
Noun phrase + of	35.0	27.5	31.8	32.3	31.6
Prepositional phrase+ of	30.6	27.4	22.8	28.1	27.2
Other prepositional phrases	14.3	6.5	9.1	10.4	10.0
Other noun phrases	6.1	9.7	9.1	7.3	8.0
Others	4.0	11.3	9.1	7.3	8.0
Passive+ prepositional phrase fragment	4.0	8.0	9.1	5.2	6.6
Anticipatory it + verb/adj	4.0	4.8	4.5	6.3	5.0
Be + noun/adjectival phrase	2.0	4.8	4.5	3.1	3.6
Total	100	100	100	100	100

Table 4 shows that noun phrases with *of* fragment (NP + of) are the most frequently used structures among the sub-corpora. In total, this group of bundles comprised 31.6% of the total number of structures used across the whole corpus. Noun phrases such as *the use of the, the presence of the, the number of the* and *a wide range of* were used in the introduction and literature review of the dissertations to assist the writers describe basic concepts and discuss various concepts and provide detailed information regarding the quantity and/or the existence of specific factors.

The second recurrent structure were the prepositional phrases with *of* fragments (PP + of). Phrases such as *in the case of, as a result of, on the basis of* and *in the process of* in total comprised 27.2% of the total number of structures across the whole corpus. It was interesting to see that there was not any significant difference between the percentage proportions of NP phrases and PP phrases in NES, NNKS and linguistics sub-corpora. Even, the NNKS writers used approximately the same percentages of NP phrases and PP phrases. This finding is in line with the findings of Rezaei et al. (2021). They found that NP phrases and PP phrases were the most frequently used structures in the abstract, introduction, and conclusion sections of applied linguistics genres. The high percentage of these structures could be associated with the preferred rhetorical patterns in the introduction and literature review sections of dissertations. Such justification could also be supported by the low percentage of passive structure bundles, shown in Table 4. Research has shown that passive bundles are frequently used in hard sciences and are widely used to discuss tables and figures (Hyland, 2008a). This might also explain the higher percentage of passive structure use in the biology dissertations in contrary to the linguistics texts. The comparison between the structures of bundles used in biology and linguistics also showed lower percentages of PP + *of* structure in biology dissertations in comparison with the linguistics ones. The two sub-corpora were, to a great extent, similar in the employment of other structures.

Functional analysis of lexical bundles

As it was discussed earlier, this study follows Hyland's (2008a, b) functional classification, according to which lexical bundles can serve research-oriented, text-oriented, and participant-oriented functions. The first category under investigation in this part is the research-oriented function. Such functions assist the writers to encode their real-world activities and experiences. Table 5 demonstrates that research-oriented bundles were the most frequently used ones across the whole corpus and in each of the groups. In a comparison between NES and NNKS writers, it can be seen that the latter made greater use of these features in their texts. Hyland (2008b) discuss that novice writers of academic texts are under pressure to show their knowledge of the research area, describing the context of the study and the physical practicalities of their study. Thus, it is plausible to say that the same reason caused the NNKS writers to use more research-oriented functions to reassure their supervisors and other expert readers of their study about their content knowledge and research skills.

Table 5. Main Bundle Functions in each Sub-corpus

Function	NES	NNKS	Biology	Linguistics	Total
Research-oriented	51.6	63.2	64.4	51.7	57.7
Text-oriented	27.0	24.8	22.6	32.2	26.7
Participant-oriented	21.4	12.0	13.0	16.1	15.6
Total (%)	100	100	100	100	100

(1) **The presence of** even a small number of peroxides in lipoproteins can significantly contribute to the subsequent oxidation in the presence of transition metal ions. (NNKS/biology)

(2) **On the basis of** the formality, IFIDs are of different types. (NNKS/linguistics)

(3) TFE stimulates **the formation of the** open promoter complex. (NES/ biology)

(4) [This thesis] covers both the nature of language change **as well as the** history of how language change has been studied up until the present day. (NES/ linguistics)

A comparison between the disciplines also demonstrated that the biology writers used more research-oriented bundles in comparison to the linguistics writers. Such a finding might be attributed to the nature of hard sciences in detailed description of the laboratory results or contexts and the environment of the study. Hyland (2008a) argues that such a tendency might be attributed to the approach of hard sciences towards research according to which they take a more empirical and research-based approach to knowledge construction.

(5) Cyclin E, essential for G1-S phase transition, **has been shown to** accumulate in differentiating human epidermal keratinocyte and in outer layers of the human epidermis. (Biology/ NNKS)

(6) Prostanoids **have been reported to** be produced in human skin and are involved **in the regulation of** growth and differentiation of the epidermis (Biology/ NES)

Text-oriented bundles were the second recurrently used functions across the whole corpus and among the four categories shown in Table 5. Text-bundles are discursive devices that enable writers to organize the texts and establish relationship between the different parts of the text. The comparison between the NES and NNKS bundles revealed that the proportion of text-oriented bundles in the dissertations written by NES writers were slightly higher than those of the NNKS writers. In other words, while the NNKS writers tend to use more research-oriented bundles, the NES writers have more inclination to use text-oriented bundles in their texts. This may suggest, besides establishing knowledge, the NES writers took the expectations of their readers into consideration and made an effort to make their texts easier to follow and more organized for their audience.

(7) **The results of the** analysis in Chapter 5 leads to further exploration of the impact of interviewer effects on negative tag realizations, presented in Chapter 6. (NES/Linguistics)

(8) **In addition to the** BCC spheroids, MSC spheroids will also be generated and co-cultured within the collagen gel. (NES/biology)

(9) The process of absorption of genistein **is shown in** Figure 2. (NNKS/biology)

Table 5 also showed that the linguistics dissertations contained more text-oriented bundles than the biology ones. Unlike the research-oriented and empiricist nature of hard sciences, soft sciences, such as linguistics in this study, are usually dominated by text-oriented bundles. This indicates that knowledge construction in soft sciences is based on the employment of discursive features which assist writers to produce plausible reasoning and arguments rather than discussing the objective findings. In such texts, writers use text-oriented bundles to create relationships between the research and the literature, to relate the ideas to the propositions and to guide the reader throughout the arguments (Hyland, 2008a).

(10) … we would expect to find evidence of it **in the form of** greatly reduced use of the traditional East Anglian variants of the two linguistic variables… . (NES/Linguistics)

(11) The data was controlled **in the sense that** they were directed to use a specific language. (NNKS/ Linguistics)

Participant-oriented bundles enable writers to directly interact with their readers through sharing their stance towards the propositions and directly engaging the readers in the discourse. Stance bundles usually

include hedging features as in clusters such as *may be due to, it is likely that,* and anticipatory-it bundles such as *it is possible that.* The engagement bundles which directly involve the readers in the discussions are basically directives such as *it should be noted that* and *it is necessary to*. Our findings showed that both groups of native and non-native writers seemed cautious in expressing their voices through using participant-oriented bundles, although, the NNKS writers used these features less frequently than the natives did. This could be traced in the rather collectivist educational principles of Kurdistan according to which individuals are usually required to avoid conveying their personal affections and expressing their epistemic judgements (Rarick, et al., 2014). This is in contrast with the individualistic principles of English writing style which encourages writers to express their stance and their uncertainties overtly (Masoumi and Lindstrom, 2009).

The employment of participant-oriented bundles also differed in biology and linguistics dissertations. Table 5 demonstrated that the linguistics texts included more participant-oriented bundles. The results revealed that in these texts stance bundles such as *it is possible to*, and *more likely to be* were widely used to assist the writers to take a tentative stance towards the presented information.

(12) While the social status of men is evaluated on "what they do", for women **it is more likely to be** on "how they appear"… . (NES/Linguistics)

(13) **It is possible to** have SVO order, though it is a marked order, as in (9). (NNKS/linguistics)

The engagement bundles, on the other hand, were the common participant-bundles used in biology dissertations. Directives used in these texts such as *it is noteworthy that*, *it is important to note that* involved the reader to cognitive acts such as noticing an issue. Moreover, biology writers used other engagement bundles such as *it is important to*, or *it should be noted that* to emphasize on the necessity or importance of the presented ideas for their readers.

(14) **It is worth noting that** the eyelid contains a specialised population of SGs called Meibomain glands. (NES/biology)

(15) **It should be noted that** the freshwater species play an important role in the world aquaculture industry as well as in annual regional finfish production. (NNKS/biology)

Conclusion and pedagogical implications

The present study investigated the distribution, structures and functions of lexical bundles used by the NES and NNKS writers across the disciplines of biology and linguistics. The results revealed that the NNKS writers used more bundles in comparison to the NES. The NNKS writers also used more research-oriented functions than their native counterparts. The considerable reliance of the NNKS writers on the use of bundles in general and the use of research-oriented bundles in specific, could perhaps reveal that the NNKS writers are less experienced than their NES ones in establishing their discourse as well as constructing and sharing their content knowledge with the experienced members of their discourse communities such as their supervisors and readers. The NES writers, on the other hand, tend to focus more on the appropriate employment of bundles and to use them for improving the organization and coherence of their texts. Moreover, both groups of writers were similar in their inclination to frequently use NP+ of and PP+ of structures in their texts. The cross-disciplinary analysis of the corpus revealed that the biology dissertations included significantly fewer bundles than the linguistics ones. Also, the biology sub-corpora included more research-oriented bundles and fewer text-oriented bundles than the linguistics texts. It was discussed that such differences might be the results of differences in the research methodologies and approaches between the two disciplines. In other words, the interpretive nature of linguistics research

and the empiricist approach of biology studies have led the writers to follow different discursive and rhetorical patterns in reporting their research steps and persuading their expert readers.

The findings of the present research should be treated with some caution. Firstly, this research mainly focused on PhD dissertations across two disciplines. Future research can be conducted on overlooked academic disciplines such as dentistry, philosophy, chemistry, and mechanical engineering. Secondly, due to some practical and technical limitations, the corpus included limited number of dissertations. Compiling larger corpus in future studies would provide more generalizable results and could provide a deeper perspective of the rhetorical preferences of the academic writers. Thirdly, this study investigated the use of lexical bundles by NNKS writers in PhD dissertations. Further research is required on the rhetorical preferences of NNKS writers across other genres and in other disciplines to gain broader understanding about the rhetorical preferences of this group of writers.

Pedagogically speaking, exploring the similarities and differences in the employment of lexical bundles between native and non-native writers across different disciplines have practical implications for EAP teachers and material developers. Including authentic text analysis tasks or focused tasks such as a series of cloze tests and gap fillings with a focus on lexical bundles used in target genres can assist the students to raise their awareness towards the structures and functions of bundles. Using extended writing tasks, also, help learners practice using bundles for various functions in different parts of the dissertation or thesis or articles. Thus, rather than providing some general academic writing textbooks for our learners, EAP learners would find the chance to specifically practice the discursive features of the genre and the discipline that they will encounter in their academic contexts. Such an approach to teaching academic writing would enable learners to successfully use various linguistic devices, in general, and lexical bundles, in specific, to improve the flow of information, coherence and fluency in their academic texts.

References

Ädel, A. and 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. http://doi.org/10.1016/j.esp.2011.08.004 *(In English)*

Amirian, Z., Ketabi, S. and Eshaghi, H. (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. *(In English)*

Appel, R. and Wood, D. (2016). Recurrent word combinations in EAP test-taker writing: Differences between high-and low-proficiency levels, *Language Assessment Quarterly*, 13 (1), 55-71. http://doi.org/10.1080/15434303.2015.1126718 *(In English)*

Ahmed, A. (2021). The employment of lexical bundles in English PhD dissertations written by English-speaking and Kurdish-speaking writers (Unpublished Master's Thesis), Near East University, Nicosia, Cyprus.

Biber, D. (2006). Stance in spoken and written university registers, *Journal of English for Academic Purposes*, 5 (2), 97-116. http://doi.org/10.1016/j.jeap.2006.05.001 *(In English)*

Biber, D., Conrad, S. and Cortes, V. (2004). If you look at…: Lexical bundles in university teaching and textbooks, *Applied linguistics*, 25 (3), 371-405. https://doi.org/10.1093/applin/25.3.371 *(In English)*

Biber, D., Johansson, S., Leech, G., Conrad, S. and Finegan, E. (1999). *Longman grammar of spoken and written English*, Pearson, London, UK. *(In English)*

Byrd, P. and 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, 31-64. *(In English)*

Candarli, D. and Jones, S. (2019). Paradigmatic influences on lexical bundles in research articles in the discipline of education, *Corpora,*

14 (2), 237-263. http://doi.org/10.3366/cor.2019.0170 *(In English)*

Cao, F. (2021). A comparative study of lexical bundles across paradigms and disciplines, *Corpora,* 16 (1), 97-128. https://doi.org/10.3366/cor.2021.0210 *(In English)*

Chen, Y. and Baker, P. (2010). Lexical bundles in L1 and L2 academic writing, *Language Learning & Technology,* 14 (2), 30–49. http://doi.org/10125/44213 *(In English)*

Cortes, V. (2004). Lexical bundles in published and student disciplinary writing: Examples from history and biology, *English for Specific Purposes,* 23 (4), 397–423. http://doi.org/10.1016/j.esp.2003.12.001 *(In English)*

Cortes, V. (2008). A comparative analysis of lexical bundles in academic history writing in English and Spanish, *Corpora,* 3 (1), 43-57. https://doi.org/10.3366/E1749503208000063 *(In English)*

Gholaminejad, R. (2021). A Comparison of two: Lexical bundles in the discourse of applied linguistics, *Atlantis Journal of the Spanish Association for Anglo-American Studies*, 43 (2), 90-109. https://doi.org/10.28914/Atlantis-2021-43.2.05 *(In English)*

Gil, N. N. and Caro, E. M. (2019). Lexical bundles in learner and expert academic writing, *Bellaterra Journal of Teaching & Learning Language & Literature,* 12 (1), 65-90. http://doi.org/10.5565/REV/JTL3.794 *(In English)*

Hernández, P. S. (2020). Lexical bundles and moves in research article abstracts of different disciplines, *Odisea: Revista de Estudios Ingleses*, 21, 71-87. http://doi.org/10.25115/odisea.v0i21.4130 *(In English)*

Hyland, K. (2008a). As can be seen: Lexical bundles and disciplinary variation, *English for Specific Purposes,* 27 (1), 4-21. http://doi.org/10.1016/j.esp.2007.06.001 *(In English)*

Hyland, K. (2008b). Academic clusters: Text patterning in published and postgraduate writing, *International Journal of Applied Linguistics,* 18 (1), 41-62. http://doi.org/10.1016/j.esp.2007.06.001 *(In English)*

Hyland, K. (2009). *Academic discourse: English in a global context*, Continuum, London, UK. *(In English)*

Jalali, Z. S. and 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. http://doi.org/10.22111/IJALS.2018.4265 *(In English)*

Jalili, H. (2013). Lexical bundles in applied linguistics: Variations across postgraduate genres, *Journal of Foreign Language Teaching and Translation Studies,* 2 (2), 1-29. http://doi.org/10.22034/EFL.2013.79199 *(In English)*

Jones, M. and Haywood, S. (2004). Facilitating the acquisition of formulaic sequences, in Schmitt, N. (ed.), *Formulaic sequences*, John Benjamins, Amsterdam, Netherlands, 269-292. *(In English)*

Kashiha, H. and Chan, S. H. (2015). A little bit about: Differences in native and non-native speakers' use of formulaic language, *Australian Journal of Linguistics,* 35 (4), 297-310. http://doi.org/10.1080/07268602.2015.1067132 *(In English)*

Masoumi, D. and Lindstrom, B. (2009). Foundations of cultural design in e-learning, *International Journal of Internet and Enterprise Management,* 6 (2), 124-142. http://doi.org/10.1504/IJIEM.2009.023926 *(In English)*

Nasrabady, P., Shirvan, M. E. and Golparvar, S. E. (2020). Exploring lexical bundles in recent published papers in the field of applied linguistics, *Journal of World Languages,* 6 (3), 175-197. http://doi.org/10.1080/21698252.2020.1797992 *(In English)*

Rarick, C., Winter, G., Barczyk, C. and Merkt, E. (2014). Iraqi and Kurdish cultural values in the semi-autonomous state of Kurdistan, *Global Journal of Management and Business Research,* 14 (3), 59-65. *(In English)*

Scott, M. and Tribble, C. (2006). *Textual patterns: Key words and corpus analysis in language education*, John Benjamins Publishing, Amsterdam, Netherlands. *(In English)*

Ucal, S. (2017). A Corpus-based study on the use of three-word lexical bundles in the academic writing by native English and Turkish non-native writers, *English Language Teaching*, 10 (12), 28-36. http://doi.org/10.5539/elt.v10n12p28 *(In English)*

Yakut, I., Yuvayapan, F. and Bada, E. (2021). Lexical bundles in L1 and L2 English doctoral dissertations, *Journal of Teaching English for Specific and Academic Purposes*, 475-493. http://doi.org/10.22190/JTESAP2103475Y *(In English)*

Yin, X. and Li, S. (2021). Lexical bundles as an intradisciplinary and interdisciplinary mark: A corpus-based study of research articles from business, biology, and applied linguistics, *Applied Corpus Linguistics,* 1 (1), 100006. http://doi.org/10.1016/j.acorp.2021.100006 *(In English)*

Zare, J. and Valipouri, L. (2021). Lexical bundles in research articles in Chemistry: A structural analysis, *Iranian Journal of English for Academic Purposes,* 10 (2), 90-105. *(In English)*

Все авторы прочитали и одобрили окончательный вариант рукописи.

All authors have read and approved the final manuscript.

Конфликты интересов: у автора нет конфликта интересов для декларации.

Conflicts of Interest: the authors have no conflict of interest to declare.

Atta A. Ahmed, MA in English Language Teaching, Assistant Lecturer, English Language Teaching Department, College of Education, Tishk International University-Sulaimani, Sulaymaniyah, Kurdistan, Iraq.

Aida Ariannejad, Ph.D. in English Language Teaching, Assistant Professor, World Peace University, Nicosia, Cyprus.